

## On ETOL Systems of Finite Index

G. ROZENBERG\*

*Department of Mathematics, University of Antwerp, UIA, B-2610 Wilrijk, Belgium*

AND

D. VERMEIR

*Department T.E.W., University of Leuven, KUL, B-3000 Leuven, Belgium*

The classical concept of finite index is investigated within the framework of ETOL systems.

### INTRODUCTION

The theory of  $L$  systems (see, e.g., Herman and Rozenberg (1975) and Rozenberg and Salomaa (1974, 1976)) constitutes today one of the most pursued fragments of formal language theory. Its relevance to the formal language theory stems from the facts that it introduced new range of problems and proof techniques and has put some of the old concepts and problems in a better perspective.

In this paper we examine the important formal language theoretic concept of a rewriting system of finite index in the framework of  $L$  systems. In particular we are concerned with the class of ETOL systems (see Rozenberg (1973b)) which form perhaps the central class in the theory of  $L$  systems. After modifying the concept of a finite index (remember that the distinction between nonterminal and terminal symbols in  $L$  systems is more subtle than in Chomsky grammars) we investigate properties of ETOL systems of finite index as well as the properties of the class of languages that they generate.

We would like to point out that the finite index restriction in ETOL systems is also biologically reasonable. It happens quite often that the development of an organism is such that things are happening in a limited number of places only. The organism may be programmed in such a way that the number of active cells do not exceed a certain threshold. Such a limitation may be also imposed by an outside controlling factor (e.g., nutrition). Since such a behavior is rather typical, we can conclude that if the concept of finite index would not have been introduced

\* Author to whom correspondence should be addressed.

already in formal language theory, certainly it would have arisen (out of biological considerations) in  $L$ -systems theory.

As to the mathematical significance of the class of languages that we investigate in this paper, namely the class of ETOL languages of finite index, we want to point out the following very interesting situation. We have investigated (see Salomaa (1973)) the effect of the finite index restriction on various classes of language-generating devices studied in the literature. It has turned out that almost all these classes (about 15, including context-free programmed grammars, matrix grammars, random context grammars, scattered grammars and ordered grammars, all under the classical finite index restriction) coincide with the class of ETOL languages of finite index. It is a rather rare situation in formal language theory and, in our opinion, it makes the research presented in this paper really important.

We assume the reader to be familiar with the rudiments of formal language theory (see, e.g., Hopcroft and Ullman (1969) or Salomaa (1973)) as well as with the rudiments of  $L$ -systems theory (see, e.g., Herman and Rozenberg (1975)).

## 1. PRELIMINARIES

The following notations will be useful in the sequel.

If  $x$  is a word, then for each positive integer  $i$ ,  $x(i)$  denotes the  $i$ th symbol of  $x$  if  $i \leq |x|$ , otherwise  $x(i) = \Delta$ .

Let  $V$  be an alphabet and let  $x$  be a word. Then  $\#_v(x)$  denotes the number of occurrences of symbols from  $V$  in  $x$ . Since in our notation we do not distinguish between a singleton and its element, we write  $\#_a(x)$  instead of  $\#_{\{a\}}(x)$  if  $a$  is a symbol.

Let  $V$  be an alphabet and let  $\Delta$  be a subset of  $V$ . Then the homomorphism  $\text{Pres}_{V,\Delta}$  ( $\text{Pres}_\Delta$  when  $V$  is understood from the context) is defined by

$$\begin{aligned} \text{Pres}_V a &= a & \text{if } a \in \Delta \\ &= \Delta & \text{if } a \in V \setminus \Delta. \end{aligned}$$

Let us now recall the notion of an ETOL system (see Rozenberg (1973b) or Herman and Rozenberg (1975)).

**DEFINITION 1.** An ETOL system is a construct  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  where  $V$  is a finite nonempty alphabet (the *alphabet of  $G$* ),

$\Sigma$  is a finite alphabet (the *target or terminal alphabet of  $G$* ),

$S \in V$  (the *axiom of  $G$* ),

$\mathcal{P}$  is a finite set each element of which (called a *table*) is a finite binary relation included in  $V \times V^*$ . It is assumed that  $(\forall P)_{\mathcal{P}} (\forall a)_V (\exists \alpha)_{V^*} (\langle a, \alpha \rangle \in P)$ .

If every  $P$  in  $\mathcal{P}$  is a subset of  $V \times V^+$  then we call  $G$  *propagating*.

If  $(\forall P)_{\mathcal{P}}(\forall a)_V(\exists! \alpha)_{V^+}(\langle a, \alpha \rangle \in P)$  then we call  $G$  *deterministic*.

If  $\#\mathcal{P} = 1$  then  $G$  is called an *EOL system*.

We use letters  $D$  and  $P$  to denote the deterministic and the propagating restriction, respectively. If  $\langle a, \alpha \rangle$  is an element of  $P$  in  $\mathcal{P}$  then we call it a *production (in  $P$ )* and we write  $a \rightarrow \alpha$  rather than  $\langle a, \alpha \rangle$ . We also write  $a \rightarrow_P \alpha$  for “ $a \rightarrow \alpha$  is in  $P$ .” The elements of  $V_N = V \setminus \Sigma$  are referred to as *nonterminals* and unless explicitly otherwise indicated we assume that  $S \in V_N$ .

DEFINITION 2. Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system.

(1) Let  $x = a_1, \dots, a_n$ , with  $a_1, \dots, a_n \in V$ , and let  $y \in V^*$ . We say that  $x$  *directly derives  $y$  (in  $G$ )*, denoted as  $x \Rightarrow_G y$ , if there is a  $P$  in  $\mathcal{P}$  such that  $y = \alpha_1 \cdots \alpha_n$  where  $a_1 \rightarrow_P \alpha_1, \dots, a_n \rightarrow_P \alpha_n$ . (In this case we also write  $x \Rightarrow_P y$ .)

(2) Let  $\stackrel{*}{\Rightarrow}_G$  be the transitive and the reflexive closure of the relation  $\Rightarrow_G$ . If  $x \stackrel{*}{\Rightarrow} y$  then we say that  $x$  *derives  $y$  in  $G$* .

(3) The *language of  $G$* , denoted as  $L(G)$ , is defined by  $L(G) = \{x \in \Sigma^* : S \stackrel{*}{\Rightarrow}_G x\}$ .

For a table  $P$  and a word  $x$  we use  $P(x)$  to define the set of all words that  $x$  directly derives “using”  $P$ . If  $G$  is deterministic and  $x \Rightarrow_P y$  then we use  $P(x)$  to denote both  $\{y\}$  and  $y$ , but this should not lead to confusion. If  $\rho$  is a sequence of (names of) tables from  $\mathcal{P}$  then we write  $x \Rightarrow^\rho y$  when  $x$  derives  $y$  using  $\rho$ . We also write  $x \Rightarrow^A y$  if  $y = x$ . We will also use  $\rho(x)$  to denote the set of all words  $y$  that can be derived from  $x$  “using”  $\rho$ . By a *derivation of  $x$  in  $G$*  we understand the precise description of how  $x$  is derived from  $S$  in  $G$ . The *trace of a derivation  $D$* , denoted as  $\text{trace}(D)$ , is the sequence of all “intermediate” words (the axiom and the final word included).

DEFINITION 3. A language  $K$  is called an ETOL (EDTOL, EPTOL) *language* if  $K = L(G)$  for an ETOL(EDTOL, EPTOL) system  $G$ .

In the sequel we use  $\mathcal{L}(X)$  to denote the class of all languages of type  $X$ . (For example,  $\mathcal{L}(\text{EPD}^*\text{TOL})$  stands for the class of all propagating and deterministic ETOL languages.) We call two ETOL systems  $G$  and  $H$  equivalent if they generate the same language (that is, if  $L(G) = L(H)$ ). We write  $L(G) = L(H)$  if  $L(G) \cup \{A\} = L(H) \cup \{A\}$ ; in other words we neglect the empty word when comparing languages generated by ETOL systems.

A feature that distinguishes nonterminal symbols from terminal symbols in Chomsky grammars is that the nonterminals are symbols which are subject to further transformation (rewriting). In ETOL systems the situation is more complicated: Both terminal and nonterminal symbols can be rewritten. To get grasp on this particular feature of a symbol we introduce now the notion of a letter active in a system.

DEFINITION 4. Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system. A letter  $a$  from  $V$  is called *active (in  $G$ )* if there exist a table  $P$  in  $\mathcal{P}$  and a word  $\alpha$  in  $V^*$  such that  $a \rightarrow_P \alpha$  and  $\alpha \neq a$ . Then  $A(G) = \{a \in V: a \text{ is active in } G\}$ .

The following result shows that one can always "organize" an ETOL system in such a way that nonterminal and active symbols coincide. Although the theorem is easy to prove it will turn out to be a very useful one.

THEOREM 1. *There exists an algorithm which given an ETOL (EDTOL) system  $G$  produces an equivalent ETOL (EDTOL) system  $H = \langle V, \mathcal{P}, S, \Sigma \rangle$  such that  $A(H) = V \setminus \Sigma$ .*

*Proof.* Let  $G = \langle Z, \mathcal{R}, S, \Sigma \rangle$ . Clearly we can assume that  $Z \setminus \Sigma \subseteq A(G)$ .

Let  $B = A(G) \cap \Sigma$  and let  $\bar{B} = \{b: b \in B\}$ . Let  $f$  be the homomorphism from  $Z^*$  into  $(Z \cup \bar{B})^*$  defined by

$$\begin{aligned} f(x) &= \bar{x} & \text{if } x \in B, \\ &= x & \text{if } x \notin B. \end{aligned}$$

Let  $\epsilon$  be a new symbol and let for every  $P$  in  $\mathcal{R}$ ,

$$f(P) = \{f(a) \rightarrow f(\alpha): a \rightarrow_P \alpha\} \cup \{a \rightarrow a: a \in \Sigma\} \cup \{\epsilon \rightarrow \epsilon\epsilon\}.$$

Let  $P_{\text{fin}}$  be a new table defined by

$$P_{\text{fin}} = \{\bar{b} \rightarrow b: b \in B\} \cup \{a \rightarrow a: a \in \Sigma\} \cup \{a \rightarrow \epsilon: a \in Z \setminus \Sigma\} \cup \{\epsilon \rightarrow \epsilon\epsilon\}.$$

Now let  $V = Z \cup \bar{B} \cup \{\epsilon\}$ ,  $\mathcal{P} = \{f(P): P \in \mathcal{R}\} \cup \{P_{\text{fin}}\}$  and let  $H = \langle V, \mathcal{P}, S, \Sigma \rangle$ . It should be clear to the reader that  $L(H) = L(G)$  and  $A(H) = V \setminus \Sigma$ . Since the construction of  $H$  from  $G$  is obviously effective and since  $H$  is deterministic if  $G$  is, the theorem holds.

If an ETOL system satisfies the conditions required from  $H$  in the statement of the above theorem then we say that it is in *active normal form*.

## 2. ETOL SYSTEMS OF FINITE INDEX

As we have already pointed out that there is a close analogy between nonterminal symbols in Chomsky grammars and active symbols in ETOL systems. Thus when introducing the notion of a finite index ETOL system, e.g., in the sense that it is used in context free grammars, one should count active rather than nonterminal symbols. This is done in the following definition.

DEFINITION 5. Let  $G$  be an ETOL system.

(1) Let  $k$  be a positive integer. We say that  $G$  is of *index*  $k$  if for every word  $x$  in  $L(G)$  there exists a derivation of  $x$  in  $G$  with the trace  $x_1, \dots, x_n$  such that, for  $1 \leq j \leq n$ ,  $\#_{A(G)}(x_j) \leq k$ .

(2) We say that  $G$  is of *finite index* if  $G$  is of index  $k$  for some  $k \geq 1$ .

DEFINITION 9. Let  $K$  be an ETOL language.

(1) Let  $k$  be a positive integer. We say that  $K$  is of *index*  $k$  if there exists an ETOL system  $G$  of index  $k$  such that  $L(G) = K$ .

(2) We say that  $K$  is of *finite index* if  $K$  is of index  $k$  for some  $k \geq 1$ .

We use  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  to denote the class of all ETOL languages of finite index  $k$  and we use  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  to denote the class of all ETOL languages of finite index. We use  $\mathcal{L}(\text{EOL})_{\text{FIN}}$ ,  $\mathcal{L}(\text{EDTOL})_{\text{FIN}(k)}$ , etc. in the same way.

The following two technical result point out simple representations of ETOL systems of finite index. Their real value shows in the proof of the Finite Index Normal Form Theorem in Section 3.

LEMMA 1. *There exists an algorithm which given an arbitrary ETOL system  $G$  of index  $k$  produces an equivalent EPTOL system  $H$  which is of index  $k$  and in active normal form.*

*Proof.* First by Theorem 1 we can obtain an ETOL system  $G'$  which is in active normal form. One easily notices that the construction given in the proof of Theorem 1 preserves the index of  $G$ , meaning that  $G'$  is also of index  $k$ .

Next we notice that the standard construction (see Rozenberg (1973a, 1973b)) to produce an EPTOL system equivalent to the given ETOL system preserves both the active normal form and the index.

Thus the lemma holds.

LEMMA 2. *There exists an algorithm which given an ETOL system  $G$  of index  $k$  produces an equivalent EPDTOL system  $H$  of index  $k$  in active normal form.*

*Proof.* Let  $G$  be an ETOL system of index  $k$ . By Lemma 1 we can assume that  $G$  is propagating and in active normal form.

Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  and let  $A(G) = V \setminus \Sigma = \{A_1, \dots, A_f\}$  with  $A_1 = S$ . Let  $B = \{A_{i,j} : i \in \{1, \dots, f\} \text{ and } j \in \{1, \dots, k\}\}$  and let  $Z = B \cup \Sigma$ . Now if  $\alpha = \alpha_0 A_{i_1} \alpha_1 A_{i_2} \alpha_2 \dots A_{i_n} \alpha_n$  with  $n \geq 0$ ,  $\alpha_0, \dots, \alpha_n \in \Sigma^*$  then we define  $\psi(\alpha) = \{\alpha_0 A_{i_1, j_1} \alpha_1 \dots A_{i_n, j_n} \alpha_n : j_1, \dots, j_n \in \{1, \dots, k\}\}$ . For  $P$  in  $\mathcal{P}$  let

$$\bar{P} = \{a \rightarrow a : a \in \Sigma\} \cup \{A_{i,j} \rightarrow \gamma : A_{i,j} \in B, A_i \rightarrow_P \alpha \text{ and } \gamma \in \psi\{\alpha\}\}$$

and let

$$\text{Det}(\bar{P}) = \{R \subseteq \bar{P} : (\forall a)_{\Sigma} (\exists ! \alpha)(a \rightarrow R \alpha)\}.$$

Now let  $H = \langle Z, \mathcal{R}, A_{11}, \Sigma \rangle$  where  $\mathcal{R} = \bigcup_{P \in \mathcal{P}} \text{Det}(\bar{P})$ . Clearly  $H$  is an EPDTOL system of index  $k$  in active normal form and  $L(H) = L(G)$ . As the construction of  $H$  from  $G$  is obviously effective the lemma holds.

As an immediate corollary of the preceding two lemmata we get the following result.

COROLLARY 1.

- (1) For every positive integer  $k$ ,  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)} = \mathcal{L}(\text{EPDTOL})_{\text{FIN}(k)}$ .
- (2)  $\mathcal{L}(\text{ETOL})_{\text{FIN}} = \mathcal{L}(\text{EPDTOL})_{\text{FIN}}$ .

### 3. ETOL SYSTEMS OF UNCONTROLLED FINITE INDEX

Among ETOL systems of finite index one can naturally distinguish these in which every successful derivation satisfies a finite index restriction. These systems are formally defined now. Note the analogous situation in the case of context free grammars of finite index versus ultralinear grammars (see e.g. Salomaa (1973)) or in the case of ETOL systems with fragmentation with outside and inside control (see Rozenberg *et al.* (1976)).

DEFINITION 7. Let  $G$  be an ETOL system.

- (1) Let  $k$  be a positive integer. We say that  $G$  is of *uncontrolled index  $k$* , if for every word  $x$  in  $L(G)$  whenever  $x_1, \dots, x_n$  is the trace of a derivation of  $x$  in  $L(G)$  then, for  $1 \leq j \leq n$ ,  $\#_{A(G)}(x_j) \leq k$ .
- (2) We say that  $G$  is of *uncontrolled finite index* if  $G$  is of uncontrolled index  $k$  for some  $k \geq 1$ .

DEFINITION 8. Let  $K$  be an ETOL language.

- (1) Let  $k$  be a positive integer. We say that  $K$  is of *uncontrolled index  $k$*  if there exists an ETOL system  $G$  of uncontrolled index  $k$  such that  $L(G) = K$ .
- (2) We say that  $K$  is of *uncontrolled finite index* if  $K$  is of uncontrolled index  $k$  for some  $k \geq 1$ .

We use  $\mathcal{L}(\text{ETOL})_{\text{FINU}(k)}$  to denote the class of all ETOL languages of uncontrolled index  $k$  and we use  $\mathcal{L}(\text{ETOL})_{\text{FINU}}$  to denote the class of all ETOL languages of uncontrolled finite index. We use  $\mathcal{L}(\text{EDTOL})_{\text{FINU}(k)}$ ,  $\mathcal{L}(\text{EPDTOL})_{\text{FINU}}$ , etc., in the similar sense.

The following result referred as the *Finite Index Normal Form Theorem* says that every ETOL system of index  $k$  can be effectively replaced by an equivalent EPDTOL system of uncontrolled index  $k$  in active normal form. This result will be very useful in proving most of the results of this paper.

**THEOREM 2.** *There exists an algorithm which given an arbitrary ETOL system  $G$  of index  $k$  produces an equivalent EPDTOL system  $H$  which is of uncontrolled index  $k$  and in active normal form.*

*Proof.* Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of index  $k$ . By Lemma 2 we can assume that  $G$  is an EPDTOL system in active normal form. Let  $V \setminus \Sigma = \{A_1, \dots, A_m\}$  with  $A_1 = S$ .

Let  $\text{Vect} = \{\langle i_1, \dots, i_m \rangle : 0 \leq i_1, \dots, i_m \leq k \text{ and } i_1 + i_2 + \dots + i_m \leq k\}$ . Let  $Z = \{B^u : B \in V \setminus \Sigma \text{ and } u \in \text{Vect}\}$  and let  $\epsilon$  be a new symbol. Let  $P \in \mathcal{P}$  and let  $u = \langle i_1, \dots, i_m \rangle \in \text{Vect}$ . Then we define

$$P(u) = \langle \#_{A_1} P(A_1^{i_1} A_2^{i_2} \dots A_m^{i_m}), \#_{A_2} P(A_1^{i_1} A_2^{i_2} \dots A_m^{i_m}), \dots, \#_{A_m} P(A_1^{i_1} \dots A_m^{i_m}) \rangle.$$

Now let  $\bar{V} = Z \cup \Sigma \cup \{\epsilon\}$ ,  $\bar{S} = S^{(1,0,0,\dots,0)}$  and let  $\bar{\mathcal{P}} = \{\bar{P} : P \in \mathcal{P}\}$  where each  $\bar{P}$  is defined as follows:

- (i) for each  $a$  in  $\Sigma$ ,  $\bar{P}(a) = \{a\}$ ,
- (ii) for each  $B^u$  in  $Z$  such that  $P(u) \in \text{Vect}$ ,

$$\begin{aligned} \bar{P}(B^u) &= \{\beta_0 C_1^{P(u)} \beta_1 \dots C_t^{P(u)} \beta_t : \beta_0, \dots, \beta_t \in \Sigma^*, C_1, \dots, C_t \in V \setminus \Sigma \text{ and} \\ &\quad P(B) = \beta_0 C_1 \beta_1 \dots C_t \beta_t\}, \end{aligned}$$

- (iii) for each  $B^u$  in  $Z$  such that  $P(u) \notin \text{Vect}$ ,  $\bar{P}(B^u) = \{\epsilon\}$ ,
- (iv)  $\bar{P}(\epsilon) = \{\epsilon\}$ .

Finally let  $H = \langle \bar{V}, \bar{\mathcal{P}}, \bar{S}, \Sigma \rangle$ .

$H$  simulates only those derivations from  $G$  that do not introduce more than  $k$  occurrences of nonterminals. It is easily done because  $G$  is deterministic and so for every string  $H$  can keep track of the total number of occurrences of nonterminals in the string ( $H$  uses elements of  $\text{Vect}$  as superscripts of nonterminals to carry this computation on). If a rewriting of a string  $x$  in  $G$  leads to a string with more than  $k$  occurrences of nonterminals then  $H$  replaces all occurrences of nonterminals in the string "simulating  $x$ " by the nonterminal  $\epsilon$ ; the so obtained string can be rewritten in  $H$  only as a string containing  $\epsilon$ 's.

These remarks should suffice to the reader to carry on the formal proof that  $L(H) = L(G)$  and that  $H$  is an EPDTOL system of uncontrolled index  $k$ . Thus the theorem holds.

It is very instructive to compare Theorem 2 with the corresponding result in the case of sequential rewriting: The class of context free languages of finite index and the class of ultralinear languages do not coincide!

If an ETOL language satisfies the conditions of the system  $H$  from the statement of the above theorem we say that it is in *Finite Index Normal Form* (abbreviated as FINF). For the sake of uniformity we will mostly assume that an ETOL system of finite index is in FINF even though some of the features of the FINF may be redundant.

As far as languages are concerned we have now the following obvious result.

COROLLARY 2. (1) *For every positive integer  $k$ ,*

$$\mathcal{L}(\text{ETOL})_{\text{FIN}(k)} = \mathcal{L}(\text{EPDTOL})_{\text{FIN}(k)} = \mathcal{L}(\text{ETOL})_{\text{FINU}(k)} = \mathcal{L}(\text{EPDTOL})_{\text{FINU}(k)}.$$

$$(2) \mathcal{L}(\text{ETOL})_{\text{FIN}} = \mathcal{L}(\text{EPDTOL})_{\text{FIN}} = \mathcal{L}(\text{ETOL})_{\text{FINU}} = \mathcal{L}(\text{EPDTOL})_{\text{FINU}}.$$

A natural question is whether one can get a normal form result which would allow us to consider only EOL systems of finite index. The answer is negative but we show that one can restrict oneself to ETOL systems of finite index with two tables only.

THEOREM 3. (1) *There exists an algorithm which for every ETOL system  $G$  of (uncontrolled) index  $k$  produces an equivalent EPTOL system  $H = \langle V, \mathcal{P}, S, \Sigma \rangle$  of (uncontrolled) index  $k$  such that  $\#\mathcal{P} = 2$ .*

(2) *There exist an ETOL language  $K$  of finite index such that for every ETOL system  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  of finite index which generates  $K$  we have  $\#\mathcal{P} \geq 2$ .*

*Proof.* (1) Let  $G = \langle Z, \mathcal{R}, U, \Sigma \rangle$  be an ETOL system of (uncontrolled) index  $k$ . We can assume that  $G$  is in FINF. Let  $\mathcal{R} = \{R_1, \dots, R_n\}$ . Let  $V = \{z_i: z \in Z \setminus \Sigma \text{ and } 1 \leq i \leq n\} \cup \Sigma$ . Let  $P_1 = \{z_i \rightarrow z_{i+1}: z \in Z \setminus \Sigma \text{ and } 1 \leq i \leq n-1\} \cup \{z_n \rightarrow z_1: z \in Z \setminus \Sigma\} \cup \{a \rightarrow a: a \in \Sigma\}$ . Let  $P_2 = \{z_i \rightarrow \alpha_{i+1}: z \in Z \setminus \Sigma, z \rightarrow_{R_i} \alpha \text{ and } 1 \leq i \leq n-1\} \cup \{z_n \rightarrow \alpha_1: z \in Z \setminus \Sigma \text{ and } z \rightarrow_{R_n} \alpha\} \cup \{a \rightarrow a: a \in \Sigma\}$  where for a word  $\alpha$  over  $Z$ ,  $\alpha_i$  denotes the word resulting from  $\alpha$  by adding subscript  $i$  to every occurrence of every nonterminal letter in  $\alpha$ .

Finally let  $H = \langle V, \mathcal{P}, S, \Sigma \rangle$  where  $\mathcal{P} = \{P_1, P_2\}$  and  $S = U_1$ . It should be obvious to the reader that

$$(i) L(G) = L(H),$$

$$(ii) H \text{ is of (uncontrolled) index } k.$$

(2) Let  $K = \{a^n b^m a^n: m \geq n \geq 1\}$ . Let  $G = \langle \{S, B, C, D, a, b\}, \{P_1, P_2\}, S, \{a, b\} \rangle$  where  $P_1 = \{S \rightarrow BC, B \rightarrow aBb, C \rightarrow Ca, D \rightarrow Db, a \rightarrow a, b \rightarrow b\}$  and  $P_2 = \{S \rightarrow S, B \rightarrow D, C \rightarrow A, D \rightarrow A, a \rightarrow a, b \rightarrow b\}$ . Clearly  $L(G) = K$ . But  $G$  is of uncontrolled index 2 and so  $K$  is an ETOL language of finite index. However it was proved in Ehrenfeucht and Rozenberg (1974b) that  $K$  is not an EOL language. Thus the second part of the theorem holds.

#### 4. AN INFINITE HIERARCHY IN $\mathcal{L}(\text{ETOL})_{\text{FIN}}$

We start this section with proving a result providing a necessary condition for an infinite language to be in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ .

First we need some definitions.



DEFINITION 9. Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of index  $k$  which is in FINF. Let  $v$  be a word in  $V^*$  such that  $\text{Pres}_{A(G)}(v) = A_1 \cdots A_p$  for some  $p \leq k$  and let  $\rho \in \mathcal{P}^*$ .

The  $\rho$ -configuration of  $v$ , denoted by  $\text{conf}(v, \rho)$  is defined by  $\text{conf}(v, \rho) = (A_1, l_1) \cdots (A_p, l_p)$  where, for  $1 \leq i \leq p$ ,  $l_i$  is defined as follows:

$$\begin{aligned} l_i &= \begin{cases} \text{MIN}\{l: \rho(1) \cdots \rho(l)(A_i) \in \Sigma^*\} & \text{if } \rho(A_i) \in \Sigma^* \text{ and} \\ = |\rho| & \text{otherwise.} \end{cases} \end{aligned}$$

We say that  $\text{conf}(v, \rho) = (A_1, l_1) \cdots (A_p, l_p)$  is *maximal* if  $p = k$  and  $l_1 = \cdots = l_p = |\rho|$ . Hence  $\text{conf}(v, \rho)$  tells us which active symbols (and in what order) are present in  $v$  and also it indicates for each active symbol to be rewritten by  $\rho$  as a terminal word how many steps it takes to do this. Then  $\text{conf}(v, \rho)$  is maximal if  $v$  contains exactly  $k$  occurrences of active symbols and none of them introduces a terminal word when we apply  $\rho$  to it.

The following result will be referred as the *Pumping Lemma*.

LEMMA 3. Let  $K \subset \Sigma^*$  be an ETOL language of index  $k$ . There exist positive integers  $d$  and  $q$  such that for every word  $w$  in  $K$  that is longer than  $d$  there exists a positive integer  $t \leq 2k$  such that  $w$  can be written in the form  $w = y_0 \alpha_1 y_1 \alpha_2 y_2 \cdots \alpha_t y_t$  with  $|\alpha_i| < q$  for  $1 \leq i \leq t$ ,  $\alpha_1 \cdots \alpha_t \neq \Lambda$  and for every positive integer  $m$ , the word  $y_0 \alpha_1^m y_1 \alpha_2^m \cdots \alpha_t^m y_t$  is in  $K$ .

*Proof.* The proof goes by induction on the index  $k$ . The case  $k = 1$  is obvious.

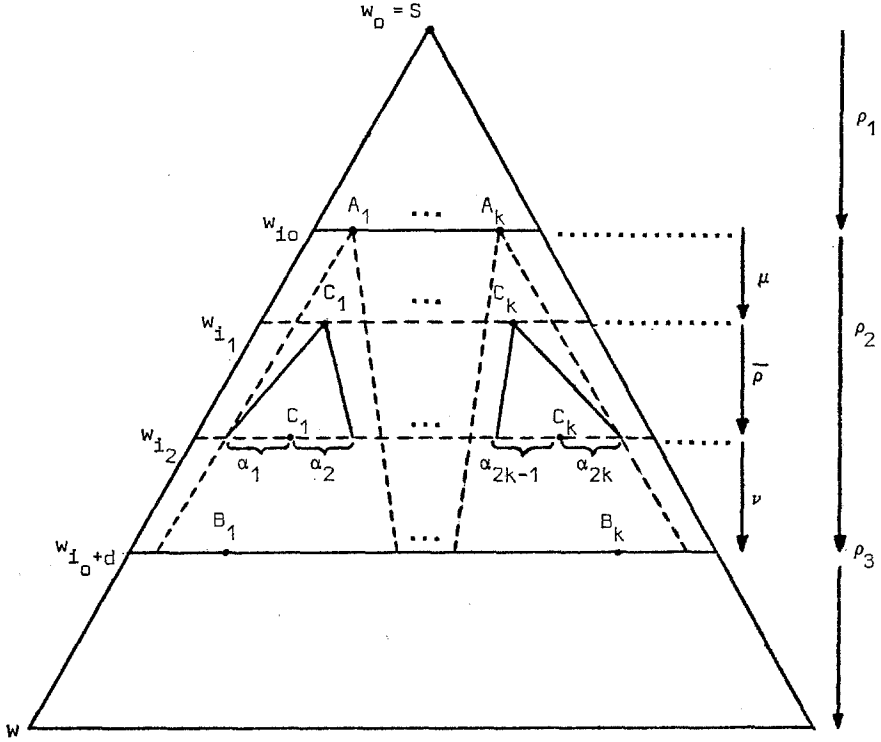
Assume that the lemma holds for index  $k - 1$  and let  $K = L(G)$  where  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  is an ETOL system of index  $k$ . We assume that  $G$  is in FINF. Let  $s = \text{MAX}\{|\alpha| : A \Rightarrow_G \alpha \text{ for some } A \in V\}$  and let  $d = (\#V)^k + 1$  and  $q = s \cdot d$ .

A derivation  $D: (S = w_0 \Rightarrow^{T_1} w_1 \Rightarrow^{T_2} \cdots \Rightarrow^{T_n} w_n = w)$  of a word  $w$  from  $L(G)$  is called *maximal* if the following conditions are satisfied.

- (i)  $n > d$ ,
- (ii) for every  $0 \leq i, j \leq n$ ,  $w_i \neq w_j$  if  $i \neq j$ , and
- (iii) there exists an integer  $i$  where  $0 \leq i < n - d$  such that  $\text{conf}(w_i, T_{i+1} \cdots T_{i+d})$  is maximal.

Let  $L_1$  be the set of all words in  $K$  that can be derived using a maximal derivation. We will first show that the lemma holds for every word in  $L_1$ .

Let  $w$  be a word in  $L_1$  and let  $D: (S = w_0 \Rightarrow^{\rho(1)} w_1 \Rightarrow^{\rho(2)} \cdots \Rightarrow^{\rho(n)} w_n = w)$  be a maximal derivation of  $w$  where  $|\rho| = n$ . Hence there exists an integer  $0 \leq i_0 < n - d$  such that  $\text{conf}(w_{i_0}, \rho_2) = (A_1, d) \cdots (A_k, d)$  where  $\rho_2 = \rho(i_0 + 1) \cdots \rho(i_0 + d)$ . Let  $\rho_1$  and  $\rho_3$  be such that  $\rho = \rho_1 \rho_2 \rho_3$ . The situation is best represented in the following diagram.



Let  $\text{Pres}_{A(G)} w_{i_0+d} = B_1 \cdots B_k$  for some  $B_1, \dots, B_k \in A(G)$ . From the definition of maximal configuration and the choice of  $d$  it follows that there exist integers  $i_0 \leq i_1 < i_2 \leq i_0 + d$  such that  $\text{Pres}_{A(G)}(w_{i_1}) = \text{Pres}_{A(G)}(w_{i_2}) = C_1 \cdots C_k$  for some  $C_1, \dots, C_k$  in  $A(G)$ . Let  $\bar{\rho} = \rho(i_1 + 1) \cdots \rho(i_2)$  and let  $\mu$  and  $\nu$  be such that  $\rho_2 = \mu \bar{\rho} \nu$ . It then follows that, for every  $1 \leq i \leq k$ ,  $\bar{\rho}(C_i) = \alpha_{2i-1} C_i \alpha_{2i}$  where  $\alpha_j \in \Sigma^*$  and  $|\alpha_j| < d \cdot s = q$  for every  $1 \leq j \leq 2k$ . Since  $w_{i_1} \neq w_{i_2}$  it also holds that  $\alpha_1 \cdots \alpha_{2k} \neq \Lambda$ . Let  $w_{i_1} = y_0 C_1 y_2 C_3 y_4 \cdots C_k y_{2k}$  and let  $\nu \rho_3(C_i) = y_{2i-1}$  for every  $1 \leq i \leq k$ . Hence  $w = y_0 \alpha_1 y_1 \alpha_2 y_2 \cdots \alpha_{2k-1} \alpha_{2k} y_{2k}$  and, obviously,  $\rho_1 \mu \bar{\rho} \nu \rho_3(S) = y_0 \alpha_1^m y_1 \alpha_2^m y_2 \cdots y_{2k-2} \alpha_{2k-1}^m y_{2k-1} \alpha_{2k}^m y_{2k}$  is in  $K$  for every  $m \geq 0$ . Hence the statement of the lemma holds for words in  $L_1$ .

Next we will show that  $L_2 = K \setminus L_1$  is an ETOL language of index  $k - 1$ . From our induction hypothesis it will then follow that the lemma holds for ETOL languages of index  $k$ , thus completing the proof.

By definition of  $L_2$ , every derivation  $D: (w_0 = S \Rightarrow^{T_1} w_1 \Rightarrow^{T_2} \cdots \Rightarrow^{T_n} w_n = w)$  of a word  $w$  from  $L_2$  has the property that, if  $n > d$  and no two words in  $\text{trace}(D)$  are equal, then  $(w_i) \# A(G) = k$  for some  $1 \leq i < n - d$  implies that at least one of the  $k$  active letters in  $w_i$  produces a terminal word in less than  $d$  steps. We will construct a new ETOL system  $H$  of index  $k - 1$  which will simulate every derivation of such type as follows. Every intermediate word in the new derivation

will contain information about the next  $d$  tables that will be applied to it. If a step in the original derivation results in a word containing  $k$  occurrences of active symbols then we know that at least one of those active symbols, say  $X$ , will produce a terminal word, say  $x$ , in less than  $d$  steps. In the new derivation such a step will be simulated by immediately substituting the occurrence  $X$  by  $x$ , thus keeping the index smaller than  $k$ . This can be done without difficulty since every word already contains information about the next  $d$  tables to be applied and thus  $x$  can easily be determined.

Here is a formal description of the construction of  $H$ .

Define a new alphabet  $V_H = \{[i, \alpha, \tau] : i < k, \tau \in \mathcal{P}^d, \alpha = \text{conf}(v, \tau) \text{ for some } v \text{ in } \bigcup_{j=1}^{k-1} A(G)^j\} \cup \{\epsilon, \$\} \cup \Sigma$  where  $\epsilon$  and  $\$$  are new symbols.

For every  $T$  and  $T'$  in  $\mathcal{P}$  we define a new table  $T_{T'}$ , as follows.

(1)  $x \rightarrow x$  is in  $T_{T'}$  for every  $x$  in  $\Sigma$ .

(2) Let  $\alpha = (A_1, t_1) \cdots (A_p, t_p)$  where  $p < k$ ,  $A_1, \dots, A_p \in A(G)$  and  $1 \leq t_i \leq d$  for  $1 \leq i \leq p$ . Let  $T\tau'$  be a word in  $\mathcal{P}^+$  such that  $|T\tau'| = d$ . For every  $1 \leq i \leq p$ , let  $T(A_i) = \beta_{i,0}B_{i,1}\beta_{i,1} \cdots B_{i_1+r_i}\beta_{i,r_i+1}$ , where  $1_1 = 1$ ,  $p_1 + r_p \leq k$ ,  $B_1, \dots, B_{p_1+r_p} \in A(G)$ , and  $\beta_{i,j} \in \Sigma^*$  for every  $1 \leq i \leq p$ ,  $0 \leq j \leq r_i + 1$ . Let  $\text{conf}(B_1 \cdots B_{p_1+r_p}, \tau'T') = (B_1, l_1) \cdots (B_{p_1+r_p}, l_{p_1+r_p}) = \beta$  be such that it is not maximal and let  $\beta'$  be the word obtained from  $\beta$  by erasing all symbols  $(B, m)$ ,  $m < d$ , from it. (Note that  $|\beta'| < k$ .) For every  $1 \leq r \leq p_1 + r_p$ , we define

$$\begin{aligned} \gamma_r &= \tau'T'(B_r) && \text{if } l_r < d \text{ and} \\ &= [r', \beta', \tau'T'] && \text{otherwise,} \end{aligned}$$

where  $r'$  is such that  $\beta'(r') = (B_r, l_r)$ . Then  $[i, \alpha, T\tau'] = \beta_{i,0}\gamma_{i,1}\beta_{i,1} \cdots \gamma_{i_1+r_i}\beta_{i,r_i+1}$  is in  $T_{T'}$  for every  $1 \leq i \leq p$ .

(3)  $X \rightarrow \epsilon$  is in  $T_{T'}$  for every  $X$  in  $V_H$ .

We also define a special initial table  $T_{\text{init}}$  as follows.

(1)  $x \rightarrow \epsilon$  for every  $x$  in  $V_H$ .

(2) Let  $\rho$  and  $\tau$  be in  $\mathcal{P}^+$  such that  $|\rho| \leq d$ ,  $|\tau| = d$  and let  $\rho(S) = v = \alpha_0 A_1 \alpha_1 \cdots A_p \alpha_p$  for some  $0 \leq p \leq k$  where  $\alpha_0, \dots, \alpha_p \in \Sigma^*$  and  $A_i \in A(G)$  for  $1 \leq i \leq p$ .

Let  $\text{conf}(v, \tau) = (A_1, l_1) \cdots (A_p, l_p) = \alpha$  be such that  $\alpha$  is not maximal. Let  $\alpha'$  be the word obtained from  $\alpha$  by erasing all symbols  $(A, m)$ ,  $m < d$ . (Note that  $|\alpha'| < k$ .) For every  $1 \leq r \leq p$  we define

$$\begin{aligned} \gamma_r &= \tau(A_r) && \text{if } l_r < d \text{ and} \\ &= [r', \alpha', \tau] && \text{otherwise where } r' \text{ is such that } \alpha'(r') = (A_r, l_r). \end{aligned}$$

Then  $\$ \rightarrow \alpha_0 \gamma_1 \alpha_1 \cdots \gamma_p \alpha_p$  is in  $T_{\text{init}}$ . Finally, let  $H = \langle V_H, \{T_T: T, T' \in \mathcal{P}\} \cup \{T_{\text{init}}\}, \$, \Sigma \rangle$ . Clearly,  $H$  is an ETOL system of index  $k - 1$ .

We leave to the reader the straightforward proof of the fact that  $L(H) = L_2$ .

By the induction hypothesis the lemma holds for  $L(H)$ . Let  $d'$  and  $q'$  be the constants of the statement of the lemma corresponding to  $L(H)$ . Let  $d'' = \text{MAX}\{d, d'\}$  and  $q'' = \text{MAX}\{q, q'\}$ . It then follows that the lemma holds for  $K$  with constants  $d''$  and  $q''$ . This completes the induction and thus the lemma holds.

The preceding result is indeed useful to provide examples of languages which are not in  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  as well as examples of languages which are not in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ . For example, the reader can easily prove now that the following holds.

**PROPOSITION 1.** *The language  $\{a^{n_1}b^{n_2}a^{n_3}b^{n_4} \cdots a^{n_{m-1}}b^{n_m}; m \geq 1 \text{ and } n_1 = n_2 = \cdots = n_m\}$  is not in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ .*

Now we prove the existence of an infinite hierarchy of classes of languages in between the class of linear context-free languages (denoted as  $\mathcal{L}(\text{LIN})$ ) and the class of ETOL languages of finite index.

**THEOREM 4.**  $\mathcal{L}(\text{LIN}) = \mathcal{L}(\text{ETOL})_{\text{FIN}(1)} \subsetneq (\text{ETOL})_{\text{FIN}(2)} \subsetneq \cdots \subsetneq \mathcal{L}(\text{ETOL})_{\text{FIN}}$

*Proof.* It is obvious that  $\mathcal{L}(\text{LIN}) = \mathcal{L}(\text{ETOL})_{\text{FIN}(1)}$ .

Now let for a positive integer  $k$ ,  $\Sigma_{2k+1}$  be a finite alphabet,  $\Sigma_{2k+1} = \{a_0, \dots, a_{2k+1}\}$  and let  $L_{k+1} = \{a_0^n a_1^n \cdots a_{2k+1}^n; n \geq 1\}$ .

Obviously  $L_{k+1} \in \mathcal{L}(\text{ETOL})_{\text{FIN}(k+1)}$ .

On the other hand the easy application of the pumping lemma yields that  $L_{k+1} \notin \mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$ .

Thus the theorem holds.

## 5. CLOSURE PROPERTIES

In this section we prove that the class of ETOL languages of finite index has a nice algebraic structure: It forms a nonprincipal full, substitution-closed, AFL in the sense of Ginsburg (1975). Moreover, for each positive integer  $k$ , the class of ETOL languages of index  $k$  forms a full semi-AFL. We also prove that  $\mathcal{L}(\text{EOL})_{\text{FIN}}$  is not even a semi-AFL.

First we notice that ETOL systems of finite index generating infinite languages possess a kind of "antisynchronization" property.

**LEMMA 4.** *Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of finite index. Let  $a$  in  $\Sigma$  be such that  $\{n: (\exists x)_{L(G)} (\#_a(x) = n)\}$  is infinite. Then the only production for  $a$  in any table of  $G$ , is  $a \rightarrow a$ .*

*Proof.* Obvious.

THEOREM 5. For every  $k \geq 1$ ,  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is a full-semi-AFL.

*Proof.* Let  $k$  be a positive integer.

(i)  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is closed under union. This can be proved by a construction identical to that in Herman and Rozenberg (1975).

(ii)  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is closed under intersection with regular languages. This can be proved using a small adaptation of the proof of the corresponding result in Herman and Rozenberg (1975).

(iii)  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is closed under substitution with regular sets.

This is proved as follows. Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of index  $k$ . We assume that  $G$  is in FINF. Let  $\psi$  be a substitution on  $\Sigma$  assigning to each  $a$  in  $\Sigma$  a regular language  $R_a$  over an alphabet  $\Delta$ .

Let  $s = \text{MAX}\{|\alpha| : a \Rightarrow_G \alpha \text{ for some } a \text{ in } V\}$ . For each word  $\alpha \in \Sigma^*$  which is not longer than  $2s$ , let  $G_\alpha = \langle N_\alpha, \Delta, P_\alpha, S_\alpha \rangle$  ( $\bar{G}_\alpha = \langle \bar{N}_\alpha, \Delta, \bar{P}_\alpha, \bar{S}_\alpha \rangle$ ) be a right- (left-) linear grammar generating  $\psi(\alpha)$  and let  $N = \bigcup_\alpha N_\alpha \cup \{e\}$ ,  $\bar{N} = \bigcup_\alpha \bar{N}_\alpha \cup \{e\}$  where  $e$  is a new symbol. We assume that all alphabets  $N_\alpha$ ,  $\bar{N}_\alpha$ , and  $V$  are mutually disjoint. Let

$$\begin{aligned} P = & \{X \rightarrow aY : X \xrightarrow{P_\alpha} aY \text{ for some } |\alpha| \leq 2s, X, Y \in N, a \in \Delta\} \\ & \cup \{X \rightarrow ae : X \xrightarrow{P_\alpha} a \text{ for some } |\alpha| \leq 2s, X \in N, a \in \Delta\} \\ & \cup \{X \rightarrow X : X \in N\} \end{aligned}$$

Similarly, let

$$\begin{aligned} \bar{P} = & \{\bar{X} \rightarrow \bar{Y}a : \bar{X} \xrightarrow{\bar{P}_\alpha} \bar{Y}a \text{ for some } |\alpha| \leq 2s, \bar{X}, \bar{Y} \in \bar{N}, a \in \Delta\} \\ & \cup \{\bar{X} \rightarrow ae : \bar{X} \xrightarrow{\bar{P}_\alpha} a \text{ for some } |\alpha| \leq 2s, \bar{X} \in \bar{N}, a \in \Delta\} \\ & \cup \{\bar{X} \rightarrow \bar{X} : \bar{X} \in \bar{N}\}. \end{aligned}$$

Define a new alphabet  $V_H = (N \times V_N \times \mathcal{P} \times N) \cup N \cup \Delta \cup \{\epsilon, \$\}$  where  $\epsilon$  and  $\$$  are new symbols. For every table  $T$  in  $\mathcal{P}$  we define a new table  $\hat{T}$  as follows.

- (1)  $a \rightarrow a$  is in  $\hat{T}$  for every  $a$  in  $\Delta$ ,
- (2)  $X \rightarrow \epsilon$  is in  $\hat{T}$  for every  $X$  in  $V_H \Sigma$
- (3) if  $A \rightarrow \alpha_0 A_1 \alpha_1 \cdots A_n \alpha_n$  is in  $T'$  for some  $T'$  in  $\mathcal{P}$ ,  $n \geq 1$ ,  $\alpha_0, \dots, \alpha_n \in \Sigma^*$ ,

and  $A, A_1, \dots, A \in V_N$ , then  $[e, A, T', e] \rightarrow X_1 \cdots X_n$  is in  $\hat{T}$  where, for  $1 < i \leq n$ ,

$$\begin{aligned} X_i &= [e, A_i, T, \bar{S}_{\alpha_i}] && \text{if } T(A_i) \notin \Sigma^*, \\ &= S_{\beta_i} && \text{otherwise, where } \beta_i = T(A_i) \alpha_i, \end{aligned}$$

and

$$\begin{aligned} X_1 &= [S_{\alpha_0}, A_1, T, \bar{S}_{\alpha_1}] && \text{if } T(A_1) \notin \Sigma^* \text{ and} \\ &= S_{\beta_1} && \text{otherwise, where } \beta_1 = \alpha_0 T(A_1) \alpha_1. \end{aligned}$$

We also define a "special" table  $T_s$  as follows.

- (1)  $X \rightarrow X$  is in  $T_s$  for every  $X$  in  $V_H$ ,
- (2) if  $X \rightarrow aY$  is in  $P$  for some  $X, Y \in N$ , and  $a = A$  or  $a \in \Delta$  and if  $\bar{X} \rightarrow \bar{Y}b$  is in  $\bar{P}$  for some  $\bar{X}, \bar{Y} \in \bar{N}$  and  $b = A$  or  $b \in \Delta$ , then  $[X, A, T, \bar{X}] \rightarrow a[Y, A, T, \bar{Y}]b$  is in  $T_s$  for each  $A$  in  $V_N$  and  $T$  in  $\mathcal{P}$  and
- (3)  $P_\alpha \subset T_s$  and  $\bar{P}_\alpha \subset T_s$  for each  $\alpha \in \Sigma^*$  with  $|\alpha| \leq 2s$ .

Finally, we need an "initial" table  $T_{\text{init}}$  which is defined by  $T_{\text{init}} = \{\$ \rightarrow [e, S, T, e]: T \in \mathcal{P}, T(S) \notin \Sigma^*\} \cup \{\$ \rightarrow S_\alpha: S \Rightarrow_G \alpha \in \Sigma^*\} \cup \{X \rightarrow x: X \in V_H\}$ . Let  $H = \langle V_H, \{\hat{T}: T \in \mathcal{P}\} \cup \{T_s, T_{\text{init}}\}, \$, \Delta \rangle$ . Clearly  $H$  is an ETOL system of index  $k$ . We leave to the reader the straightforward proof of the fact that  $L(H) = \psi(L(G))$ .

The following is an immediate consequence of (iii).

- (iv)  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is closed under homomorphism.
- (v)  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  is closed under inverse homomorphism.

This follows from (i), (ii), (iii) and the well-known result (see, e.g., Ginsburg (1975)) that a class of languages, closed under union, with regular sets, intersection with regular sets, and regular substitution is also closed under inverse homomorphism. The theorem then follows from (i) to (v).

**THEOREM 6.**  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is a substitution closed full AFL.

*Proof.* (i)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under union. This can be proved by a construction identical to that in Herman and Rozenberg (1975).

(ii)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under catenation. This can be proved by a construction identical to that in Herman and Rozenberg (1975).

(iii)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under intersection with regular languages. This can be proved using a small adaptation of the proof of the corresponding result in Herman and Rozenberg (1975).

(iv)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under substitution. This is proved as follows.

Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of finite index. We assume that  $G$  is in FINF. Let  $\psi$  be a substitution on  $\Sigma$  assigning to each  $a$  in  $\Sigma$  an ETOL language of finite index  $L_a$ . Let for each  $a$  in  $\Sigma$ ,  $G_a = \langle V_a, \mathcal{P}_a, S_a, \Sigma \rangle$  be an ETOL system of finite index such that  $L(G_a) = L_a$ . We assume that each  $G_a$  is in FINF.

Let  $Z = \bigcup_{a \in \Sigma} A(G_a)$ . We can clearly assume that  $Z \cap A(G) = \emptyset$  and that all alphabets  $A(G_a)$  are pairwise disjoint. Let for each  $a$  in  $\Sigma$ ,  $S'_a$  be a new symbol and let  $Z' = Z \cup \{S'_a : a \in \Sigma\}$ . Let  $\epsilon$  be a new symbol and let  $\overline{A(G)} = \{\bar{a} : a \in A(G)\}$ . Let  $\gamma$  be a homomorphism on  $V^*$  defined by:

$$\begin{aligned} \gamma(a) &= \bar{a} & \text{if } a \in A(G), \\ &= S'_a & \text{if } a \in \Sigma. \end{aligned}$$

For  $P$  in  $\mathcal{P}$ ,  $\bar{P} = \{a \rightarrow \gamma(a) : a \rightarrow_P \alpha\} \cup \{\bar{a} \rightarrow \bar{a} : a \in A(G)\} \cup \{a \rightarrow \epsilon : a \in Z' \cup \{\epsilon\}\}$ . Let  $P_0 = \{\bar{a} \rightarrow a : a \in A(G)\} \cup \{a \rightarrow a : a \in \Sigma\} \cup \{a \rightarrow \epsilon : a \in Z' \cup \{\epsilon\} \cup A(G)\}$ . Let  $a \in \Sigma$  and let  $P \in \mathcal{P}_a$ . Then  $\bar{P} = P \cup \{S'_a \rightarrow S_a, S'_a \rightarrow S_a\} \cup X_a \cup \{b \rightarrow b : b \notin V_a \cup \{S'_a\}\}$ , where

$$\begin{aligned} X_a &= \emptyset & \text{if } A \notin L_a, \\ &= \{S'_a \rightarrow A\} & \text{if } A \in L_a. \end{aligned}$$

Let  $\bar{V} = V \cup \overline{A(G)} \cup Z' \cup \{\epsilon\}$ . Let  $\mathcal{R} = \{\bar{P} : P \in \mathcal{P}\} \cup \{\bar{P} : P \in \bigcup_{a \in \Sigma} \mathcal{P}_a\} \cup \{P_0\}$ . Finally let  $H = \langle \bar{V}, \mathcal{R}, S, \Sigma \rangle$ .

It should be clear to the reader that:

(1)  $H$  is of finite index. One easily notices that each word in  $L(H)$  can be derived by iterating the following "macro-step": First apply a table from  $\{\bar{P} : P \in \mathcal{P}\}$ ; then apply a sequence of tables from  $\{\bar{P} : P \in \bigcup_{a \in \Sigma} \mathcal{P}_a\}$  so as to dispose of all occurrences of symbols from  $Z'$  and finally apply  $P_0$  which changes all elements from  $\overline{A(G)}$  into their unbarred counterparts. Thus the index of  $H$  does not exceed  $m.n.l$ , where  $m$  is the index of  $G$ ,  $n$  is the maximal index among  $G_a$ 's and  $l$  is the maximal length of the right-hand side of a production in  $G$ .

$$(2) L(H) = \psi(L(G)).$$

Thus (iv) holds.

But then, in particular,  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under arbitrary homomorphisms.

(v)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under inverse homomorphism.

This follows from (i), (iii), (iv) and a standard result in AFL theory (see, e.g., Lemma 9.4 in Hopcroft and Ullman (1969)) which says that if a class of languages is closed under ( $\mathcal{A}$ -free) substitution, ( $k$ -limited) erasing, and union and intersection with regular sets then it is closed under an inverse homomorphism.

(vi)  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is closed under the cross operator.

This is proved as follows: Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of uncontrolled finite index. Let  $U$  be a new symbol. Let  $P_0$  be a new table defined by  $P_0 = \{U \rightarrow US, U \rightarrow S\} \cup \{a \rightarrow a: a \in V\}$ . Let for each  $P$  in  $\mathcal{P}$ ,  $P' = P \cup \{U \rightarrow U\}$ . Let  $H = \langle V \cup \{U\}, \{P_0\} \cup \{P': P \in \mathcal{P}\}, U, \Sigma \rangle$ . It should be obvious to the reader that  $H$  is of finite index (the index of  $H$  does not exceed  $m + 1$ , where  $m$  is the index of  $G$ ) and  $L(H) = (L(G))^+$ .

From (i) to (vi) it follows that  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is a substitution closed full AFL. As a corollary of Theorems 5 and 6 we obtain the following.

COROLLARY 3.  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is non-full-principal.

*Proof.* From Theorem 5 it follows that  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  is the union of an infinite chain of full semi-AFL's. Hence, by a result from Ginsburg (1975), it is a non-full-principal full semi-AFL. Since a (full) AFL is (full) principal if and only if it is a (full) principal semi-AFL (corollary from Theorem 5.4.1 in Ginsburg (1975)), it follows that the corollary holds.

Now we will turn to closure properties of  $\mathcal{L}(\text{EOL})_{\text{FIN}}$ . First we show a specific language to be not in EOL language of finite index.

LEMMA 5.  $L_0 = \{a^n b^n c^n: n \geq 1\} \notin \mathcal{L}(\text{EOL})_{\text{FIN}}$ .

*Proof.* We will prove the lemma by demonstrating that if  $G$  is an EOL system of finite index such that  $L_0 \subseteq L(G)$  then  $L(G) \setminus L_0 \neq \emptyset$ . As a matter of fact we can consider propagating EOL systems only, because again by the standard construction (see Rozenberg and Salomaa (1976)) one easily shows that  $\mathcal{L}(\text{EOL})_{\text{FIN}} = \mathcal{L}(\text{EPOL})_{\text{FIN}}$ .

Thus let  $G = \langle V, P, S, \Sigma \rangle$  be an EPOL system of index  $k$  such that  $L_0 \subseteq L(G)$ . Let  $s$  be the length of the longest right-hand side of a production in  $P$  and let  $m$  be the number of active symbols in  $G$ . Let  $N = s^{m+1}$ .

Notice that, by Lemma 4, the only productions for  $a$ ,  $b$ , and  $c$  in  $P$  are  $a \rightarrow a$ ,  $b \rightarrow b$ ,  $c \rightarrow c$ .

Let  $w = a^n b^n c^n$  with  $n > N$ . Let  $D = (S = x_0, x_1, \dots, x_t = w)$  be a derivation of  $w$  in  $G$  with  $T$  being its derivation tree. Furthermore we assume that  $D$  is "reduced" in the sense that there do not exist a node  $x$  in  $T$  and integers  $i, j$  with  $i < j$  such that the label of  $x$  on the level  $i$  is an active symbol and the only descendant of  $x$  on the level  $j$  has the same label. Also  $x_{t-1} \neq x_t$ . Clearly such a derivation  $D$  exists.

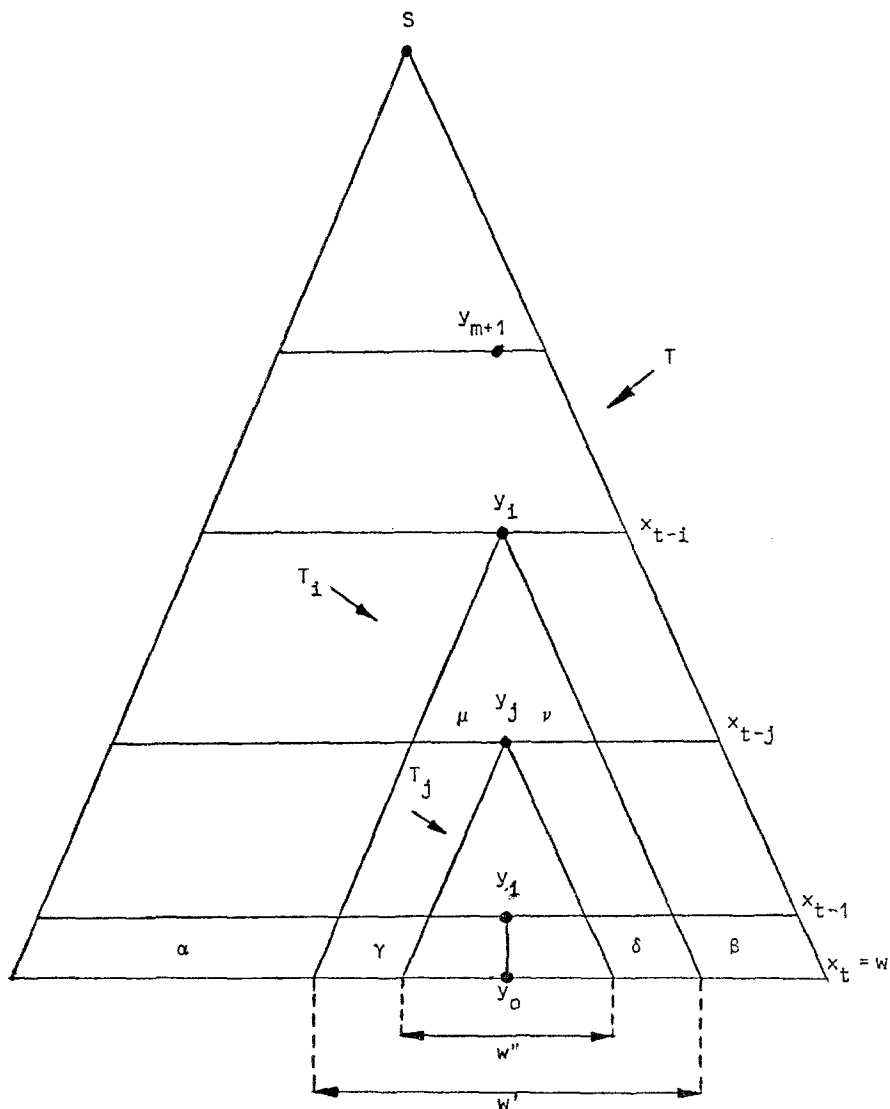
Thus there exists a node  $y_0$  on the level  $t$ , whose direct ancestor (father)  $y_1$  (on the level  $t - 1$ ) is an active symbol. Let us consider now the chain

$$l(y_1), l(y_2), \dots, l(y_{m+1})$$

of direct ancestors. (We use  $l(z)$  to denote the label of the node  $z$ ). Clearly all of them are active symbols and, since  $m = \#A(G)$ , there must be a repetition. That is for  $i \neq j$  we have  $l(y_i) = l(y_j) = A$ .



The situation is the best represented by the following picture.



Let  $\mu A\nu$  be the contribution of  $y_i$  on the level  $t - j$ . Since  $D$  is reduced,  $\mu\nu \neq A$ . Since  $n > N$ ,  $w'$  (which is the contribution of  $y_i$  on the level  $t$ ) is shorter than  $n$ . Thus  $w''$  (which is the contribution of  $y_j$  on the level  $t$ ) is a proper subword of  $w'$ .

Obviously, exchanging in  $T$  the subtree  $T_i$  by the subtree  $T_j$  yields new derivation tree  $T'$  of a word  $\alpha w''\beta$ . Thus  $\alpha w''\beta$  is in  $L(G)$ .

To complete the proof it suffices to show that  $\alpha w''\beta$  is not in  $L_0$ .

Since  $l(y_0) \in \{a, b, c\}$  and since all three cases are symmetric let us assume that  $l(y_0) = b$ . But  $|w'| < n$  and so  $\#_a(w') = 0$  or  $\#_c(w') = 0$ . As these cases are symmetric let us assume that  $\#_a(w') = 0$ . Hence  $\#_a(\gamma\delta) = 0$  and consequently

$$\#_a(\alpha w''\delta) = \#_a(w) \quad (1)$$

Since  $\gamma\delta \neq A$ , we have

$$\#_{\{b,c\}}(\alpha w''\delta) < \#_{\{b,c\}}(w) \quad (2)$$

But (1) and (2) imply that  $\alpha w''\delta \notin L_0$ , which completes the proof of the lemma.

As a direct consequence of Lemma 5 we get the following result.

**COROLLARY 4.** *There exist EOL languages that are not in  $\mathcal{L}(\text{EOL})_{\text{FIN}}$  but are in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ .*

*Proof.* Take  $L_0$  from Lemma 5. It is a well-known example of an EOL language. But the EDTOL system of index 2  $G = \langle \{S, A, B, a, b, c\}, \{P_1, P_2\}, S, \{a, b, c\} \rangle$  with  $P_1 = \{S \rightarrow AB, A \rightarrow aAb, B \rightarrow Bc, a \rightarrow a, b \rightarrow b, c \rightarrow c\}$  and  $P_2 = \{S \rightarrow abc, A \rightarrow ab, B \rightarrow c, a \rightarrow a, b \rightarrow b, c \rightarrow c\}$  generates  $L_0$ .

Now we will prove that  $\mathcal{L}(\text{EOL})_{\text{FIN}}$  is not an AFL.

**THEOREM 7.**  *$\mathcal{L}(\text{EOL})_{\text{FIN}}$  is not an AFL.*

*Proof.* We will show that  $\mathcal{L}(\text{EOL})_{\text{FIN}}$  is not closed even with respect to letter-to-letter homomorphism (coding).

Let  $K = \{da^n eb^n fc^n : n \geq 0\}$ . Obviously  $K \in \mathcal{L}(\text{EOL})_{\text{FIN}}$ . Let  $\varphi$  be the coding from  $\{a, b, c, d, e, f\}$  into  $\{a, b, c\}$  defined by  $\varphi(d) = \varphi(a) = a$ ,  $\varphi(e) = \varphi(b) = b$  and  $\varphi(f) = \varphi(c) = c$ . Then  $\varphi(K) = \{a^n b^n c^n : n \geq 1\} = L_0$  which by Lemma 5 is not in  $\mathcal{L}(\text{EOL})_{\text{FIN}}$ .

## 6. SOME BASIC DECISION PROBLEMS

There are some basic decision problems that one ought to consider when introducing a new class of languages. In particular if we say that a system from the class  $X$  is a "type- $Y$ " system if it satisfies property  $P$  we should ask ourselves whether the property  $P$  is decidable in  $X$ .

First we show that the property of having index  $k$  (having a finite index) is not decidable in the class of ETOL systems.

**THEOREM 8.** (1) *There is no algorithm which given an arbitrary ETOL system  $G$  and a positive integer  $k$  decides whether or not  $G$  is of index  $k$ .*

(2) *There is no algorithm which given an arbitrary ETOL system  $G$  decides whether or not  $G$  is of finite index.*

*Proof.* As usual we will encode the Post Correspondence Problem (see, e.g., Hopcroft and Ullman (1969)).

(1) Let  $k$  be a positive integer and let  $Z = \langle \alpha_1, \dots, \alpha_n \rangle$ ,  $W = \langle \beta_1, \dots, \beta_n \rangle$  be an instance of the Post Correspondence Problem over an alphabet  $\Sigma$ .

Let  $G(k, Z, W) = \langle V, P, U, \Sigma \cup \{\epsilon, *\} \rangle$  be an ETOL system such that (for a word  $z$ ,  $\text{mir}(z)$  denotes the mirror image of  $z$ ):

(i)  $A(G(k, Z, W)) = \{U, S, A, B, C, D, E, F, M\}$ .

(ii)  $P$  consists of the following productions

$$\begin{aligned}
 U &\rightarrow (S*)^k, \\
 U &\rightarrow E(*E)^{k-1}F, \\
 S &\rightarrow S, \\
 S &\rightarrow xAx, && \text{for every } x \text{ in } \Sigma, \\
 S &\rightarrow xBy, && \text{for every } x, y \text{ in } \Sigma \text{ such that } x \neq y, \\
 A &\rightarrow xAx, && \text{for every } x \text{ in } \Sigma, \\
 A &\rightarrow xBy, && \text{for every } x, y \text{ in } \Sigma \text{ such that } x \neq y, \\
 A &\rightarrow xC, && \text{for every } x \text{ in } \Sigma, \\
 A &\rightarrow Dx, && \text{for every } x \text{ in } \Sigma, \\
 B &\rightarrow xB, && \text{for every } x \text{ in } \Sigma, \\
 B &\rightarrow Bx, && \text{for every } x \text{ in } \Sigma, \\
 B &\rightarrow \epsilon, \\
 C &\rightarrow xC, && \text{for every } x \text{ in } \Sigma, \\
 C &\rightarrow \epsilon, \\
 D &\rightarrow Dx, && \text{for every } x \text{ in } \Sigma, \\
 D &\rightarrow \epsilon, \\
 E &\rightarrow E, \\
 E &\rightarrow \alpha_i M \text{ mir}(\beta_i), && \text{for every } i \text{ in } \{1, \dots, n\}, \\
 M &\rightarrow \alpha_i M \text{ mir}(\beta_i), && \text{for every } i \text{ in } \{1, \dots, n\}, \\
 M &\rightarrow \epsilon, \\
 F &\rightarrow F, \\
 F &\rightarrow *, \\
 x &\rightarrow x, && \text{for every } x \text{ in } \Sigma \cup \{*, \epsilon\}
 \end{aligned}$$

It should be clear to the reader that every word in  $L(G(k, Z, W))$  can be derived in such a way that the first production used is  $U \rightarrow (S*)^k$  if and only if the given instance  $Z, W$  of the postcorrespondence problem has no solution. But if the first production used in a derivation is  $U \rightarrow E(*E)^{k-1}F$  then already the first word used contains  $(k+1)$  cocurrences of active symbols. Consequently,  $G(k, Z, W)$  is of index  $k$  if and only if the given instance  $Z, W$  of the Post Correspondence Problem has no solution. (Clearly, in

this case  $L(G(k, Z, W)) = \{\alpha_1 * \alpha_2 * \dots * \alpha_k * : \alpha_i \in \{\beta \notin \text{mir}(\gamma) : \beta, \gamma \in \Sigma^+, \beta \neq \gamma\}$  for  $1 \leq i \leq k\}$ .) Since the Post Correspondence Problem is undecidable, the first part of the theorem holds.

(2) The proof of this part of the theorem is quite similar to the proof of the first part of the theorem.

Let  $Z = \langle \alpha_1, \dots, \alpha_n \rangle$ ,  $W = \langle \beta_1, \dots, \beta_n \rangle$  be an instance of the Post Correspondence Problem over an alphabet  $\Sigma$ .

Let  $G(Z, W) = \langle V, \{P_0, P_1\}, T, \Sigma \cup \{\epsilon, *\} \rangle$  be an ET0L system such that

(i)  $A(G(Z, W)) = \{T, U, S, A, B, C, D, E, M, \bar{E}\}$ ,

(ii)  $P_0$  consists of the following production

$$\begin{aligned} T &\rightarrow U, \\ T &\rightarrow \bar{E}^*, \\ U &\rightarrow US^*, \\ U &\rightarrow A, \\ \bar{E} &\rightarrow \bar{E} * \bar{E}, \\ x &\rightarrow x, \quad \text{for every } x \text{ in } V \setminus \{T, U, \bar{E}\}, \end{aligned}$$

(iii)  $P_1$  consists of the following productions

$$\begin{aligned} S &\rightarrow S, \\ S &\rightarrow xAx, \quad \text{for every } x \text{ in } \Sigma, \\ S &\rightarrow xBy, \quad \text{for every } x, y \text{ in } \Sigma \text{ such that } x \neq y, \\ A &\rightarrow xAx, \quad \text{for every } x \text{ in } \Sigma, \\ A &\rightarrow xBy, \quad \text{for every } x, y \text{ in } \Sigma \text{ such that } x \neq y, \\ A &\rightarrow xC, \quad \text{for every } x \text{ in } \Sigma, \\ A &\rightarrow Dx, \quad \text{for every } x \text{ in } \Sigma, \\ B &\rightarrow xB, \quad \text{for every } x \text{ in } \Sigma, \\ B &\rightarrow Bx, \quad \text{for every } x \text{ in } \Sigma, \\ B &\rightarrow \epsilon, \\ C &\rightarrow xC, \quad \text{for every } x \text{ in } \Sigma, \\ C &\rightarrow \epsilon, \\ D &\rightarrow Dx, \quad \text{for every } x \text{ in } \Sigma, \\ D &\rightarrow \epsilon, \\ \bar{E} &\rightarrow E, \\ E &\rightarrow \alpha_i M \text{ mir}(\beta_i), \quad \text{for every } i \text{ in } \{1, \dots, n\}, \\ M &\rightarrow \alpha_i M \text{ mir}(\beta_i), \quad \text{for every } i \text{ in } \{1, \dots, n\}, \\ M &\rightarrow \epsilon, \\ x &\rightarrow x, \quad \text{for every } x \text{ in } \Sigma \cup \{\epsilon, *\} \cup \{T, U\}. \end{aligned}$$

The reader can easily see that  $G(Z, W)$  is of finite index if and only if the given instance  $Z, W$  of the Post Correspondence Problem has no solution. Thus the second part of the theorem holds.

The above result is a "negative" one; but we are used already to live with such results in formal language theory. However the situation in our case is much better than usually. We know already that if we are interested in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  we can restrict ourselves to ETOL systems of uncontrolled finite index. We will prove now that the property of having an uncontrolled finite index (uncontrolled index  $k$ ) is decidable in the class of ETOL systems.

**THEOREM 9.** (1) *There is an algorithm which given an arbitrary ETOL system  $G$  and a positive integer  $k$  decides whether or not  $G$  is of uncontrolled index  $k$ .*

(2) *There is an algorithm which given an arbitrary ETOL system  $G$  decides whether or not  $G$  is of uncontrolled finite index.*

*Proof.* Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system. Let us assume that  $G$  is in active normal form.

Let for each  $P$  in  $\mathcal{R}$ ,  $\bar{P}$  be the table defined as follows,  $\bar{P} = \{a \rightarrow \text{Pres}_{A(G)}(\alpha) : a \in V \setminus \Sigma \text{ and } \alpha \rightarrow_P \alpha\}$ .

Now let  $H = \langle V \setminus \Sigma, \{\bar{P} : P \in \mathcal{P}\}, S, V \setminus \Sigma \rangle$ .

It is well known (see Ginsburg and Rozenberg (1975)) that the set of all sequences of tables leading from the axiom to a terminal word in an ETOL system is regular and it can be effectively found for a given ETOL system. Let this control set for  $G$  be  $\phi_G$ . Clearly the set  $\text{Pref}(\phi_G)$  of all nonempty prefixes of words in  $\phi_G$  is also regular and can be effectively given.

Now let us consider the ETOL system  $H$  with the control set  $\text{Pref}(\phi_G)$ . It is well known (see Ginsburg and Rozenberg (1975)) that one can effectively construct an ETOL system  $I$  such that  $L(I)$  equals the set of all words in  $L(H)$  derived under the control  $\text{Pref}(\phi_G)$ .

(1) Let  $k$  be a positive integer. It is clear that  $G$  is of uncontrolled index  $k$  if and only if  $I$  does not contain a word of length longer than  $k$ . But, obviously, it is decidable whether an arbitrary ETOL system generates a word longer than a fixed constant. Thus the first part of the theorem holds.

(2) It is clear that  $G$  is of uncontrolled finite index if and only if the language of  $I$  is finite. But it is decidable whether an arbitrary ETOL system generates a finite language (see, e.g., Rozenberg (1973b)) and thus the second part of the theorem holds.

Since already  $\mathcal{L}(\text{ETOL})_{\text{FIN}(1)}$  is identical to  $\mathcal{L}(\text{LIN})$  we have the following obvious result.

**THEOREM 10.** (1) *There is no algorithm which given an arbitrary pair of ETOL systems  $G$  and  $H$ , both of (uncontrolled) index 1, decides whether or not  $L(G) = L(H)$ .*

(2) *There is no algorithm which given an arbitrary pair of ETOL systems  $G$  and  $H$ , both of (uncontrolled) finite index, decides whether or not  $L(G) = L(H)$ .*

7. A  $k$ -UNIVERSAL ETOL SYSTEM

In this section we show that for every positive integer  $k$  the class  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  has a rather interesting representation. There exists a single EPDTOL system  $G_k$  of index  $k$  such that for every language  $K$  from  $\mathcal{L}(\text{ETOL})_{\text{FIN}(k)}$  one can find a suitable control set  $\phi_K$  (over tables from  $G_k$ ) such that  $K$  is the language generated by  $G_k$  under the control  $\phi_K$ .

First we note that the known result (see Ginsburg and Rozenberg (1975)) that adding regular control to ETOL systems does not get us languages beyond  $\mathcal{L}(\text{ETOL})$  holds also for ETOL systems of finite index. As the proof can be made totally analogous to the prove of the corresponding result in Ginsburg and Rozenberg (1975) we leave it to the reader. (We use the standard notation  $L_\phi(G)$  to denote the language generated by  $G$  under the control  $\phi$ ).

**THEOREM 11.** *There exists an algorithm which given any ETOL system  $G$  of uncontrolled index  $k$  and a finite automaton  $\Gamma$  produces an ETOL system  $H$  of uncontrolled index  $k$  such that  $L(H) = L_{L(\Gamma)}(G)$ .*

Now we will prove the existence of a  $k$ -universal ETOL system over a fixed terminal alphabet  $\Sigma$  meaning an ETOL system  $G$  with the property that whenever an ETOL system  $H$  of index  $k$  (over  $\Sigma$ ) is given one can construct a regular set  $\phi_H$  such that  $L_{\phi_H}(G) = L(H)$ . Note that our  $k$ -universal ETOL system is itself an ETOL system of uncontrolled index  $k$ .

**THEOREM 12.** *Let  $\Sigma$  be a finite alphabet. There exists an algorithm which for every positive integer  $k$  constructs an EPDTOL system  $G_k$  of uncontrolled index  $k$  such that there exists an algorithm which given an arbitrary ETOL system  $H$  of index  $k$  (over  $\Sigma$ ) constructs a finite automaton  $\Gamma_H$  such that  $L_{L(\Gamma_H)}(G_k) = L(H)$ .*

*Proof.* (Sketch.)

Let  $Z_1 = \{N_t^{(1)}: 1 \leq t \leq k\}$ ,  $\bar{Z}_1 = \{\bar{N}_t^{(1)}: 1 \leq t \leq k\}$ ,  $Z_2 = \{N_t^{(i)}: 1 \leq t \leq k \text{ and } 2 \leq i \leq k\}$ ,  $\bar{Z}_2 = \{\bar{N}_t^{(i)}: 1 \leq t \leq k \text{ and } 2 \leq i \leq k\}$ , and  $Z = \Sigma \cup Z_1 \cup \bar{Z}_1 \cup Z_2 \cup \bar{Z}_2$ . Let

- (1) for each  $a$  in  $\Sigma$  and  $1 \leq t \leq k$

$$T_{t,a,1} = \{N_t^{(1)} \rightarrow a\} \cup \{X \rightarrow X: X \in Z \setminus \{N_t^{(1)}\}\},$$

$$\bar{T}_{t,a,1} = \{\bar{N}_t^{(1)} \rightarrow a\} \cup \{X \rightarrow X: X \in Z \setminus \{\bar{N}_t^{(1)}\}\},$$

- (2) for each  $a$  in  $\Sigma$ ,  $1 \leq t \leq k$ , and  $2 \leq i \leq k$

$$T_{t,a,i} = \{N_t^{(i)} \rightarrow a\} \cup \{N_r^{(j)} \rightarrow N_r^{(j-1)}: 1 \leq r \leq t, 2 \leq j \leq k, \text{ and } N_r^{(j)} \neq N_t^{(i)}\}$$

$$\cup \{\bar{N}_r^{(j)} \rightarrow \bar{N}_r^{(j-1)}: 1 \leq r \leq t \text{ and } 2 \leq j \leq k\}$$

$$\cup \{X \rightarrow X: X \in \Sigma \cup Z_1 \cup \bar{Z}_1\},$$

$$\begin{aligned}\bar{T}_{t,a,t} = & \{\bar{N}_t \rightarrow a\} \cup \{\bar{N}_r^{(j)} \rightarrow \bar{N}_r^{(j-1)}: 1 \leq r \leq t, 2 \leq j \leq k, \text{ and } \bar{N}_r^{(j)} \neq \bar{N}_t^{(i)}\} \\ & \cup \{N_r^{(j)} \rightarrow N_r^{(j-1)}: 1 \leq r \leq t \text{ and } 2 \leq j \leq k\} \\ & \cup \{X \rightarrow X: X \in \Sigma \cup Z_1 \cup Z_1\},\end{aligned}$$

(3) for each  $a$  in  $\Sigma$ ,  $1 \leq t \leq k$ , and  $1 \leq i \leq k$

$$\begin{aligned}T_{t,a,i,t} &= \{N_t^{(i)} \rightarrow aN_t^{(i)}\} \cup \{X \rightarrow X: X \in Z \setminus \{N_t^{(i)}\}\}, \\ \bar{T}_{t,a,i,t} &= \{\bar{N}_t^{(i)} \rightarrow a\bar{N}_t^{(i)}\} \cup \{X \rightarrow X: X \in Z \setminus \{\bar{N}_t^{(i)}\}\}, \\ T_{t,a,i,t,\text{right}} &= \{N_t^{(i)} \rightarrow N_t^{(i)}a\} \cup \{X \rightarrow X: X \in Z \setminus \{N_t^{(i)}\}\}, \\ \bar{T}_{t,a,i,t,\text{right}} &= \{\bar{N}_t^{(i)} \rightarrow \bar{N}_t^{(i)}a\} \cup \{X \rightarrow X: X \in Z \setminus \{\bar{N}_t^{(i)}\}\},\end{aligned}$$

(4) for each  $1 \leq t \leq k$ ,  $1 \leq u \leq k$ , and  $1 \leq i \leq k$

$$\begin{aligned}T_{t,i,u} &= \{N_t^{(i)} \rightarrow \bar{N}_u^{(i)}\} \cup \{N_r^{(j)} \rightarrow N_r^{(j)}: 1 \leq r \leq k, 1 \leq j \leq k, \text{ and } N_r^{(j)} \neq N_t^{(i)}\} \\ & \cup \{X \rightarrow X: X \in Z \setminus \{Z_1 \cup Z_2\}\}, \\ \bar{T}_{t,i,u} &= \{\bar{N}_t^{(i)} \rightarrow \bar{N}_u^{(i)}\} \cup \{X \rightarrow X: X \in Z \setminus \{\bar{N}_t^{(i)}\}\},\end{aligned}$$

(5) for each  $1 \leq t \leq k$ ,  $1 \leq u \leq k$ , and  $1 \leq i \leq k-1$

$$\begin{aligned}\hat{T}_{t,i,u} &= \{\bar{N}_t^{(i)} \rightarrow \bar{N}_t^{(i+1)}\bar{N}_u^{(i+1)}\} \\ & \cup \{\bar{N}_r^{(j)} \rightarrow \bar{N}_r^{(j+1)}: 1 \leq r \leq k, 1 \leq j \leq k-1, \text{ and } \bar{N}_r^{(j)} \neq \bar{N}_t^{(i)}\} \\ & \cup \{N_r^{(j)} \rightarrow N_r^{(j+1)}: 1 \leq r \leq k \text{ and } 1 \leq j \leq k-1\} \\ & \cup \{\bar{N}_r^{(k)} \rightarrow \bar{N}_r^{(k)}\} \cup \{N_r^{(k)} \rightarrow N_r^{(k)}: 1 \leq r \leq k\} \\ & \cup \{X \rightarrow X: X \in \Sigma\}\end{aligned}$$

(6)  $T_{\text{fin}} = \{\bar{N}_t^{(j)} \rightarrow N_t^{(j)}: 1 \leq j \leq k \text{ and } 1 \leq t \leq k\} \cup \{X \rightarrow X: X \in Z \setminus Z_2\}.$

Finally let  $G_k = \langle Z, \mathcal{R}, N_1^{(1)}, \Sigma \rangle$  where  $\mathcal{R}$  is the collection of tables defined above.

First of all the reader should note that in a derivation in  $G_k$  if  $x$  is a word appearing in it then all occurrences of nonterminals in  $x$  have the same superscript ( $j$ ) which equals the total number of occurrences of active symbols in  $x$ . In this way it is assured that no word appearing in a derivation in  $x$  has more than  $k$  occurrences of active symbols. So  $G_k$  is indeed an EPDTOL system of uncontrolled index  $k$ .

Now we want to show that for an arbitrary ETOL system  $H$  of index  $k$  one can effectively construct a finite automation  $\Gamma_H$  such that the set of all words over  $\Sigma$  derived in  $G_k$  under the regular control  $L(\Gamma_H)$  equals  $L(H)$ .

The formal definition of the automation  $\Gamma_H$  would be rather lengthy (and unreadable) so instead we try to describe how it works.

Let us assume that  $H = \langle V, \mathcal{P}, U, \Sigma \rangle$  is in FINF with  $\mathcal{P} = \{R_1, \dots, R_m\}$  and  $A(H) = \{A_1, \dots, A_n\}$ .

$G_k$  under the control  $L(\Gamma_H)$  simulates  $H$  as follows.

Suppose that the string  $x = x_0 A_1 x_1 \cdots A_t x_t$  with  $1 \leq t \leq k$ ,  $x_0, \dots, x_t \in \Sigma^*$ , and  $A_1, \dots, A_t \in A(G)$  is derived in  $H$ . Then  $G_k$  under the control  $L(\Gamma_H)$  will derive a string of the form  $\tilde{x} = x_0 N_{i_1}^{(t)} x_1 \cdots N_{i_t}^{(t)} x_t$  and  $\Gamma_H$  will get into the state remembering that  $N_{i_1}^{(t)}$  stands for  $A_1$ ,  $N_{i_2}^{(t)}$  stands for  $A_2$ , etc. Now if productions  $A_1 \rightarrow \alpha_1, \dots, A_t \rightarrow \alpha_t$  are applied in  $H$  then  $G_k$  will simulate these applications in de facto sequential way. It will attempt to start at any  $N_{i_j}^{(t)}$  and rewrite it by  $\alpha_j$  changed in such a way that nonterminals from  $\alpha_j$  will be substituted by "markers" from  $\bar{Z}_1 \cup \bar{Z}_2$  and this will be done by a sequence of smaller steps available in tables of  $G_k$ . The difficulty is that the index  $k$  cannot be exceeded and so  $\Gamma_k$  may interrupt the sequence simulating  $\alpha_j$  and go to another sequence so as to get the superscript in markers  $N_r^{(j)}$  or  $\bar{N}_r^{(j)}$  to be smaller than  $k$ . At this moment another marker staying for a nonterminal can be again introduced and the sequence simulating rewriting  $A_j \rightarrow \alpha_j$  can be continued further. Such interrupts are always possible because  $H$  is of uncontrolled index  $k$ .

For example, let  $t = k = 4$ ,  $x_0 = aa$ ,  $x_1 = x_2 = A$ ,  $x_3 = b$ ,  $x_4 = bc$ ,  $A_1 = A_3 = B$ ,  $A_2 = A_4 = C$ , and let productions  $A_1 \rightarrow aBC$ ,  $A_2 \rightarrow c$ ,  $A_3 \rightarrow ab$ ,  $A_4 \rightarrow BCb$  be applied.

Let  $\tilde{x} = aa N_2^{(4)} N_4^{(4)} N_1^{(4)} b N_3^{(4)} bc$  be a string derived in  $G_k$  (controlled by  $\Gamma_H$ ) corresponding to  $x$ . A possible sequence of tables from  $G_k$  simulating the rewriting of  $x$  in  $H$  may be like this

- (1)  $T_{2,a,4,2}$  is applied to  $x$  yielding

$$y_1 = aaa N_2^{(4)} N_4^{(4)} N_1^{(4)} b N_3^{(4)} bc,$$

- (2)  $T_{2,4,1}$  is applied to  $y_1$  yielding

$$y_2 = aaa \bar{N}_1^{(4)} N_4^{(4)} N_1^{(4)} b N_3^{(4)} bc,$$

and  $\Gamma_H$  is in a state remembering that  $\bar{N}_1^{(4)}$  stands for  $B$ .

(3) Now the next step in simulating the rewriting of  $A_1$  by  $aBC$  should be the application of table  $\hat{T}_{1,4,2}$ . This however cannot be done because the superscript (4) would have to be increased to (5) which is impossible. (That would correspond to exceeding the index of  $H$ .) Thus the sequence simulating  $A_1 \rightarrow aBC$  is interrupted ( $\Gamma_H$  remembers the phase of interruption) and  $\Gamma_H$  attempts to apply a rule decreasing the superscript to (3).

- (4) To this aim the table  $T_{4,c,4}$  is applied yielding

$$y_3 = aaa \bar{N}_1^{(3)} c N_1^{(3)} b N_3^{(3)} bc.$$

This completes the simulation of the production  $A_2 \rightarrow c$ .



(5) Now  $\Gamma_H$  "returns" to complete the simulation of  $A_1 \rightarrow aBC$ . To this aim  $\hat{T}_{1,3,2}$  is applied yielding

$$y_4 = aaa\bar{N}_1^{(4)}\bar{N}_2^{(4)}cN_1^{(4)}bN_3^{(4)}bc$$

and  $\Gamma_H$  remembers that  $\bar{N}_1^{(4)}$  stands for  $B$  and  $\bar{N}_2^{(4)}$  stands for  $C$ . This completes the simulation of the production  $A_1 \rightarrow aBC$ .

(6) To start the simulation of  $A_3 \rightarrow ab$  the table  $T_{1,a,4,1}$  is applied yielding

$$y_5 = aaa\bar{N}_1^{(4)}\bar{N}_2^{(4)}caN_1^{(4)}bN_3^{(4)}bc.$$

(7) The table  $T_{1,b,3}$  is applied yielding

$$y_6 = aaa\bar{N}_1^{(3)}\bar{N}_2^{(3)}cabbN_3^{(3)}bc$$

and this completes the simulation of the rule  $A_3 \rightarrow ab$ .

(8) To start the simulation of  $A_4 = BCb$  the table  $T_{3,3,3}$  is applied yielding

$$y_7 = aaa\bar{N}_1^{(3)}\bar{N}_2^{(3)}cabb\bar{N}_3^{(3)}bc$$

and  $\Gamma_H$  remembers that  $\bar{N}_3^{(3)}$  stand for  $B$ .

(9) The table  $\hat{T}_{3,3,4}$  is applied yielding

$$y_8 = aaa\bar{N}_1^{(4)}\bar{N}_2^{(4)}cabb\bar{N}_3^{(4)}\bar{N}_4^{(4)}bc$$

and  $\Gamma_H$  remembers that  $\bar{N}_4^{(4)}$  stands for  $C$ .

(10) Now the table  $\bar{T}_{4,b,4,4, \text{right}}$  is applied yielding

$$y_9 = aaa\bar{N}_1^{(4)}\bar{N}_2^{(4)}cabb\bar{N}_3^{(4)}\bar{N}_4^{(4)}bbc$$

and this completes the simulation of the rule  $A_3 \rightarrow ab$ .

(11) Now the table  $T_{\text{fin}}$  is applied yielding

$$y_{10} = aaaN_1^{(4)}N_2^{(4)}cabbN_3^{(4)}N_4^{(4)}bbc$$

completing the simulation of rewriting  $x$  by the rules  $A_1 \rightarrow aBC$ ,  $A_2 \rightarrow c$ ,  $A_3 \rightarrow ab$  and  $A_4 \rightarrow BCb$ .

Thus while in  $H$  the string  $aaaBCcabbBCbbc$  was obtained, in  $G_n$  controlled by  $\Gamma_H$  the string  $aaaN_1^{(4)}N_2^{(4)}cabbN_3^{(4)}N_4^{(4)}bbc$  was obtained and  $\Gamma_H$  remembers that  $N_1^{(4)}$  stands for  $B$ ,  $N_2^{(4)}$  stands for  $C$ ,  $N_3^{(4)}$  stands for  $B$ , and  $N_4^{(4)}$  stands for  $C$ . The initial state of  $\Gamma_H$  is the state remembering that  $N_1^{(1)}$  stands for  $U$  and the final state is the one which remembers that there are no more nonterminals (markers) in the string under rewriting. (To this state one arrives after applying

the table which is either of the form  $T_{t,a,1}$  or of the form  $\bar{T}_{t,a,1}$ ,  $1 \leq t \leq k$ ,  $a \in \Sigma$ .)

We hope that the above description suffices to the reader to write down (if necessary) the (horrible) formal description of  $\Gamma_H$ .

The reader should be also convinced that  $L_{L(\Gamma_H)}(G_k) = L(H)$ , which ends the proof of the theorem.

The following result, which is the main result of this section, follows now easily from the previous two theorem.

**COROLLARY 5.** *Let  $\Sigma$  be a finite alphabet. For every positive integer  $k$  there exists an EPDTOL system  $G_k$  of uncontrolled index  $k$  such that  $\{L_C(G_k): C \text{ is a regular set}\}$  is identical with the family of ETOL languages of index  $k$  over  $\Sigma$ .*

## 8. INTERRELATIONS OF $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ WITH SOME FAMILIES OF LANGUAGES

In this section we investigate the position of  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$  among several ("naturally related") classes of languages.

First we prove that as far as the sets of Parikh vectors are concerned the class of ETOL systems of finite index is exactly as powerful as the class of right-linear grammars. This result is obviously useful in providing examples of languages not in  $\mathcal{L}(\text{ETOL})_{\text{FIN}}$ . (In what follows, for a language  $K$ ,  $\text{Par}(K)$  denotes the set of Parikh vectors of words in  $K$  and for a class of languages  $X$ ,  $\text{PAR}(X) = \{\text{Par}(K): K \in X\}$ .)

**THEOREM 13.**  $\text{PAR}(\mathcal{L}(\text{ETOL})_{\text{FIN}}) = \text{PAR}(\mathcal{L}(\text{CF})) = \text{PAR}(\mathcal{L}(\text{REG}))$ .

*Proof.* (i) It is well known (see Parikh (1966)) that  $\text{PAR}(\mathcal{L}(\text{CF})) = \text{PAR}(\mathcal{L}(\text{REG}))$ .

(ii) Since  $\mathcal{L}(\text{REG}) \subseteq \mathcal{L}(\text{ETOL})_{\text{FIN}}$ ,  $\text{PAR}(\mathcal{L}(\text{REG})) \subseteq \text{PAR}(\mathcal{L}(\text{ETOL})_{\text{FIN}})$ .

We shall prove now that:

(iii) There exists an algorithm which given an arbitrary ETOL system  $G$  of uncontrolled finite index produces a right-linear grammar  $H$  such that  $\text{Par}(L(G)) = \text{Par}(L(H))$ .

This is proved as follows.

Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system of controlled index  $k$ . We assume that  $G$  is in FINF. Let  $H = \langle Z_N, \Sigma, P, U \rangle$  be a right-linear grammar such that:

- (1)  $Z_N = \{[B_1 \cdots B_r]: 1 \leq r \leq k \text{ and } B_1, \dots, B_r \in V \setminus \Sigma\}$ ,
- (2)  $U = [S]$ .
- (3)  $P$  consists of the following productions:

(3.1)  $[B_1 \cdots B_r] \rightarrow \alpha$  for  $\alpha$  in  $\Sigma^*$  if and only if there exists a table  $T$  in  $\mathcal{P}$  such that  $B_1 \cdots B_r \Rightarrow_T \alpha$ ,

(3.2)  $[B_1 \cdots B_r] \rightarrow \alpha[B_1' \cdots B_s']$  for  $\alpha$  in  $\Sigma^*$  if and only if there exists a table  $T$  in  $\mathcal{P}$  such that  $B_1 \cdots B_r \Rightarrow_T \beta$ ,  $\text{Pres}_E(\beta) = \alpha$ ,

$$\text{Pres}_{V \setminus E}(\beta) = B_1' \cdots B_s'.$$

It should be clear to the reader that  $\text{Par}(L(G)) = \text{Par}(L(H))$ .

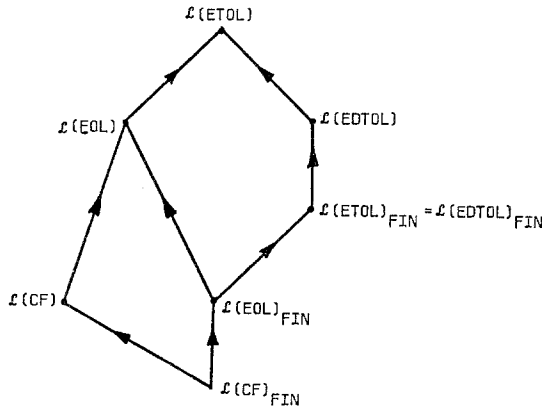
Now the theorem follows from (i) to (iii).

In the same way as context-free languages are naturally represented by EOL systems (see, e.g., Herman and Rozenberg (1975)), context-free languages of finite index are naturally represented by EOL systems of finite index. As the proof of the following result is standard, we leave it to the reader.

**THEOREM 14.** *A language  $K$  is a context-free language of finite index if and only if it can be generated by an EOL system  $G$  of finite index such that  $a \rightarrow a$  is a production in  $G$  for each terminal symbol  $a$ .*

Now we can prove the main result of this section. (In what follows,  $\mathcal{L}(\text{CF})$  denotes the family of context-free languages, and  $\mathcal{L}(\text{CF})_{\text{FIN}}$  denotes the family of context-free languages of finite index.)

**THEOREM 15.** *The following diagram holds:*



A solid line denotes strict inclusion in the direction indicated. If two families  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are not connected by a path following the arrows in the diagram, then the families are incomparable but not disjoint.

*Proof.* (i) It is well known (see Herman and Rozenberg (1975)) that

$$(i.1) \quad \mathcal{L}(EOL) \not\subseteq \mathcal{L}(ETOL),$$

$$(i.2) \quad \mathcal{L}(EDTOL) \not\subseteq \mathcal{L}(ETOL),$$

$$(i.3) \quad \mathcal{L}(CF) \not\subseteq \mathcal{L}(EOL),$$

$$(i.4) \quad \mathcal{L}(EOL) \text{ and } \mathcal{L}(EDTOL) \text{ are incomparable but not disjoint.}$$

(ii) By definition  $\mathcal{L}(EOL)_{FIN} \subseteq \mathcal{L}(EOL)$ . By Theorem 13,  $\{a^{2^n}: n \geq 0\}$  is not in  $\mathcal{L}(EOL)_{FIN}$  but it is obviously in  $\mathcal{L}(EOL)$ . Thus  $\mathcal{L}(EOL)_{FIN} \not\subseteq \mathcal{L}(EOL)$ .

(iii) By Theorem 14,  $\mathcal{L}(CF)_{FIN} \subseteq \mathcal{L}(EOL)_{FIN}$ . On the other hand  $\{ea^n db^n fc^n: n \geq 0\}$  is clearly in  $\mathcal{L}(EOL)_{FIN}$  while it is not in  $\mathcal{L}(CF)_{FIN}$ .

(iv) By Theorems 6 and 7,  $\mathcal{L}(EOL)_{FIN} \not\subseteq \mathcal{L}(ETOL)_{FIN}$  and by Corollary 2,  $\mathcal{L}(ETOL)_{FIN} = \mathcal{L}(EDTOL)_{FIN}$ .

(v) By definition  $\mathcal{L}(EDTOL)_{FIN} \subseteq \mathcal{L}(EDTOL)$  but  $\mathcal{L}(EDTOL)$  contains  $\{a^{2^n}: n \geq 1\}$  which by Theorem 13 is not in  $\mathcal{L}(EDTOL)_{FIN}$ .

(vi) The language  $\{a^n b^m a^n: m \geq n \geq 1\}$  is clearly in  $\mathcal{L}(ETOL)_{FIN}$  while it is not in  $\mathcal{L}(EOL)$  (see Ehrenfeucht and Rozenberg (1974b)). On the other hand  $\{a^{2^n}: n \geq 1\}$  is in  $\mathcal{L}(EOL)$  while it is not in  $\mathcal{L}(ETOL)_{FIN}$  (see Theorem 13).

(vii)  $\mathcal{L}(CF)$  contains languages which are not in  $\mathcal{L}(EDTOL)$  (see Ehrenfeucht and Rozenberg (1974a)) and  $\{ea^n db^n fc^n: n \geq 0\}$  is in  $\mathcal{L}(EOL)_{FIN}$  while it is not in  $\mathcal{L}(CF)$ . Now the diagram follows from (i) to (vii).

## 9. DISCUSSION

In this paper we have investigated the properties of ETOL systems and languages of finite index. It turned out that the notion of an uncontrolled finite index ETOL system may be more "natural" than the notion of a finite index ETOL system. First of all, all languages in  $\mathcal{L}(ETOL)_{FIN}$  can be generated by ETOL systems of uncontrolled finite index; it is very instructive to compare this situation with the sequential case: The class of context free languages of finite index and the class of ultralinear languages do not coincide. Moreover, while it is decidable whether an arbitrary ETOL system is of uncontrolled finite index, it is not decidable whether an arbitrary ETOL system is of finite index. We have also shown that increasing the index of an ETOL system leads to a greater generative power. This yielded us an infinite hierarchy of classes of languages between  $\mathcal{L}(LIN)$  and  $\mathcal{L}(ETOL)_{FIN}$ . It was rather interesting to see that, for a given  $k$ , all ETOL systems of index  $k$  can be represented by a single ETOL system varying regular control.

It turned out that the family of ETOL languages of finite index is a rather "decent" family: It forms an AFL.

One can say that our definition of a finite index was a “static” one. For a given derivation we are counting the number of active symbols, i.e., symbols which *can* be replaced in a system by something else. However, they do not have to be rewritten (in a given derivation) by something else. (It would be so only if the system would be in active normal form.)

To contrast this one can count, for a given derivation, the number of symbols actually rewritten in a derivation step by something else than themselves. This leads to the notion of an ETOL system of *dynamic finite index* (we leave its formal definition to the reader). However, it turns out that the class of these systems is equivalent in their language generating power to the class of ETOL systems.

**THEOREM 16.** *A language is an ETOL language if and only if it can be generated by an ETOL system of dynamic index 1.*

*Proof.* Clearly every language generated by an ETOL system of dynamic finite index is an ETOL language.

To prove the converse we proceed as follows.

Let  $G = \langle V, \mathcal{P}, S, \Sigma \rangle$  be an ETOL system and let us assume that  $G$  is in active normal form.

Let  $\overline{A}(G) = \{\bar{a} : a \in A(G)\}$  and  $V_p = \{[P], \overline{[P]} : P \in \mathcal{P}\}$ . We define a new alphabet  $Z = V \cup \overline{A}(G) \cup V_p \cup \{U, \mathfrak{c}\}$  where  $U$  and  $\mathfrak{c}$  are new symbols. Let

$$P_{\text{init}} = \{U \rightarrow [P]S, U \rightarrow \overline{[P]}S : P \in \mathcal{P}\} \cup \{a \rightarrow a : a \in Z \setminus \{U\}\}.$$

Let

$$P_{\text{term}} = \{[P] \rightarrow \Lambda, \overline{[P]} \rightarrow \Lambda : P \in \mathcal{P}\} \cup \{a \rightarrow a : a \in \Sigma\} \\ \cup \{a \rightarrow \mathfrak{c} : a \in Z \setminus (\Sigma \cup V_p)\}.$$

Let

$$P_{\text{switch}} = \{[P] \rightarrow \overline{[T]} : P, T \in \mathcal{P}\} \cup \{a \rightarrow a : a \in \Sigma \cup \overline{A}(G)\} \\ \cup \{a \rightarrow \mathfrak{c} : a \in Z \setminus (\Sigma \cup \overline{A}(G))\}.$$

Let

$$\overline{P}_{\text{switch}} = \{\overline{[P]} \rightarrow [T] : P, T \in \mathcal{P}\} \cup \{a \rightarrow a : a \in \Sigma \cup A(G)\} \\ \cup \{a \rightarrow \mathfrak{c} : a \in Z \setminus (\Sigma \cup A(G))\}.$$

For every  $P$  in  $\mathcal{P}$  let

$$P' = \{a \rightarrow \bar{\alpha} : a \xrightarrow{p} \alpha \text{ and } \alpha \in A(G)\} \cup \{X \rightarrow \mathfrak{c} : X \in V_p \setminus \{[P]\}\} \\ \cup \{a \rightarrow a : a \in (Z \setminus V_p) \cup \{[P]\}\}, \text{ where for a word } \alpha, \bar{\alpha}$$

denotes the word resulting from  $\alpha$  by barring all occurrences of elements from  $A(G)$  in  $\alpha$ ,

$$P'' = \{\bar{a} \rightarrow \alpha: a \xrightarrow{p} \alpha \text{ and } a \in A(G)\} \cup \{X \rightarrow \epsilon: X \in V_p \setminus \{\overline{[P]}\}\} \\ \cup \{a \rightarrow a: a \in (Z \setminus V_p) \cup \{\overline{[P]}\}\}.$$

Finally, let  $H = \langle Z, \mathcal{R}, V, \Sigma \rangle$  where  $\mathcal{R} = \{P_{\text{init}}, P_{\text{term}}, P_{\text{switch}}, \bar{P}_{\text{switch}}\} \cup \{P', P'': P \in \mathcal{P}\}$ . It should be clear to the reader that  $L(H) = L(G)$  and that indeed  $H$  is an ETOL system of dynamic index 1.

Thus the theorem holds.

It is instructive to notice that there is a *limited* analogy between ETOL systems of dynamic index 1 and context-free grammars. As in context-free grammars every word in the language of an ETOL system of dynamic index 1 can be obtained by a derivation such that at most one symbol is "really rewritten" in a single derivation step of this derivation. However if a derivation is not successful and leading to a word  $x$  there may be no way of deriving this word in  $G$  using a derivation of dynamic index 1.

RECEIVED: July 1, 1976

## REFERENCES

- EHRENFEUCHT, A., AND ROZENBERG, G. (1974a), On some context free languages that are not deterministic ETOL languages, Dept. of Computer Science, University of Colorado at Boulder, Technical report #CU-CS-048-74.
- EHRENFEUCHT, A., AND ROZENBERG, G. (1974b), On  $\Theta$ -definite EOL languages, Dept. of Computer Science, University of Colorado at Boulder, Technical Report #CU-CS-046-74.
- GINSBURG, S. (1975), "Algebraic and Automata-Theoretic Properties of Formal Languages," North-Holland, Amsterdam/Oxford.
- GINSBURG, S., AND ROZENBERG, G. (1975), TOL schemes and control sets, *Inform. Contr.* **27**, 109-125.
- HERMAN, G. T., AND ROZENBERG, G. (1975), "Developmental Systems and Languages," North-Holland, Amsterdam/Oxford.
- HOPCROFT, J., AND ULLMAN, J. (1969), "Formal Languages and Their Relation to Automata," Addison-Wesley, Reading, Mass.
- PARIKH, R. (1966), On context-free languages, *J. Assoc. Comput. Mach.* **13**, 570-581.
- ROZENBERG, G. (1973a), TOL systems and languages, *Inform. Contr.* **23**, 357-381.
- ROZENBERG, G. (1973b), Extensions of tabled OL systems and languages, *Internat. J. Comput. Inf. Sci.* **2**, 311-334.
- ROZENBERG, G., RUOHONEN, K., AND SALOMAA, A. (1976), Developmental systems with fragmentation, *Internat. J. Comput. Math.* **5**, 177-191.
- ROZENBERG, G., AND SALOMAA, A. (EDS.) (1974), "L Systems," Lecture Notes in Computer Science 15, Springer-Verlag, Heidelberg.

- ROZENBERG, G., AND SALOMAA, A. (1976), The mathematical theory of  $L$  systems, in "Advances in Information Systems Science 3" (J. Tou, Ed.), Vol. 6, Plenum, New York.
- ROZENBERG, G., AND VERMEIR, D., On the effect of the finite index restriction on several families of grammars, submitted for publication.
- SALOMAA, A. (1973), "Formal Languages," Academic Press, London/New York.